AFRL-RY-WP-TR-2010-1230

ON THE SATISFACTION OF MODULUS AND AMBIGUITY FUNCTION CONSTRAINTS IN RADAR WAVEFORM OPTIMIZATION FOR DETECTION

Lee K. Patton

Wright State University

JUNE 2010
Final Report

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the USAF 88th Air Base Wing (88 ABW) Public Affairs Office (PAO) and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RY-WP-TR-2010-1230 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

*//Signature//

JOHNATHAN SCANLAN
Program Manager
Radar Signal Processing Branch
RF Sensor Technology Division

//Signature//

WILLIAM BALDYGO, Chief
RF Sensor Technology Division
Sensors Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

*Disseminated copies will show "//Signature//" stamped or typed above the signature blocks.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
June 2010	Final	01 June 2008 – 01 April 2010

4. TITLE AND SUBTITLE
ON THE SATISFACTION OF MODULUS AND AMBIGUITY FUNCTION CONSTRAINTS IN RADAR WAVEFORM OPTIMIZATION FOR DETECTION

5a. CONTRACT NUMBER
In-house

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62204F

6. AUTHOR(S)
Lee K. Patton

5d. PROJECT NUMBER
7622

5e. TASK NUMBER
11

5f. WORK UNIT NUMBER
76221114

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Wright State University

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYRT

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TR-2010-1230

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
PAO Case Number: 88ABW 09-2925; Clearance Date: 26 Jun 2009. This report contains color.

This Report is a Ph.D. dissertation written by a Wright State University (Dayton, OH) doctoral candidate in 2009.

14. ABSTRACT
We consider the design of radar systems that are capable of using knowledge of their interference environment to dynamically design transmit waveforms that afford optimum signal-to-interference-plus-noise ratio while satisfying modulus and ambiguity function constraints. We begin by establishing the inextricable nature of modulus constraints in the waveform optimization problem. We then extend the state of the art in waveform optimization to accommodate these constraints. This is done by solving a secondary optimization problem using the method of alternating projections. We demonstrate that this approach can be a computationally efficient alternative to dynamic programming methods. We then consider the multiple-target detection problem, which is the basis for introducing ambiguity function constraints into the waveform design process. We formulate the waveform optimization problem for several receiver architectures, and solve these problems using sequential quadratic programming and interior point methods. Finally, we address the need for a more computationally tractable approach by considering a number of suboptimal formulations. This includes a novel formulation based on a parameterization of nonlinear frequency modulation.

15. SUBJECT TERMS
radar, waveform design

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Johnathan Scanlan
Unclassified	Unclassified	Unclassified	SAR	168	19b. TELEPHONE NUMBER *(Include Area Code)* N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

Contents

List of Figures

List of Tables

Acknowledgment

As is the case with any sufficiently arduous task, this dissertation would not have been possible had it not been for the contributions of many individuals. First of all, I would like to thank my adviser, Dr. Brian Rigling, who secured the initial funding that made it possible for me to leave industry and devote myself to this work. It was Dr. Rigling who introduced me to the topic that would become the subject of this dissertation, namely, waveform optimization. Dr. Rigling is an excellent teacher and a study in efficiency. I feel lucky to have had him as a mentor. I must also thank Dr. Mike Bryant of the Air Force Research Laboratory (AFRL) for bringing me into the fold at AFRL. Employment at AFRL has provided me with both financial support and access to great researchers. Both factors have been integral to my success, and I have to thank Dr. Bryant for that. My gratitude must also be extended to my supervisor at AFRL, Mr. Bill Baldygo, who has a keen understanding of how researchers do their work and how innovation happens. Mr. Baldygo is creating an incubator of innovation at AFRL, and I feel lucky to be a part of it. I must also thank my many teachers at Wright State and the Air Force Institute of Technology. Too numerous to name here, each has played a crucial role in my intellectual development. I would also like to acknowledge the significant influence that my colleague, Mr. Ed Zelnio, has had on my personal and professional education. It would be difficult to overstate the degree to which my education has been enhanced by our lengthy conversations in pursuit of a Theory of Everything. Thanks to Mr. Zelnio, I have seen a panorama of ideas that would have been otherwise obscured by my own esoteric research interests. Finally, I must express an unending debt of gratitude to my wife Kristen, who has with great patience and grace borne the many extra burdens that have arisen from my far too frequent absence. I am not sure how I would have completed this work in my professional life had I not had the rich personal life that she provides.

To my wife Kristen, and my children, Elise and baby Patton.

Mathematical Notation and Convention

A description of the notation most commonly used throughout this dissertation is provided in Table 1. Additionally, we adopt the following conventions: The distance between two vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$d(\mathbf{x}, \mathbf{y}) := ||\mathbf{x} - \mathbf{y}|| \tag{1}$$

whereas the distance between a vector $\mathbf{x}$ and a set S is defined as

$$d(\mathbf{x}, S) := \inf_{\mathbf{s} \in S} d(\mathbf{x}, \mathbf{s}). \tag{2}$$

The projection of a vector $\mathbf{x}$ onto a set S is defined as

$$\Pi_S(\mathbf{x}) := \{\mathbf{s} \in S : d(\mathbf{x}, \mathbf{s}) = d(\mathbf{x}, S)\}. \tag{3}$$

In general, Π_S is set-valued and may be empty. Finally, in Ch. 3, a tilde will be used to denote membership in the codomain of an operator (e.g., $\tilde{\mathbf{x}}$), whereas a hat will denote membership in the range of an operator (e.g., $\widehat{\mathbf{x}}$).

x	scalar variable
$\mathbf{x}$	vector variable
x_n	n^{th} element of vector $\mathbf{x}$
$\mathbf{X}$	matrix
$[\mathbf{X}]_{m,n}$ or $x_{m,n}$	element in m^{th} row and n^{th} column of matrix $\mathbf{X}$
j	imaginary unit ($j^2 = -1$)
$\lvert\cdot\rvert$	absolute value (modulus) operator
$\angle\cdot$	angle (argument) of a complex value, we adopt $\angle 0 := 0$
$\mathrm{Re}\{\cdot\}$	real part of a complex value
$\mathrm{Im}\{\cdot\}$	imaginary part of a complex value
$(\cdot)^*$	conjugate of a complex number
$(\cdot)^T$	transpose of a vector
$(\cdot)^H$	conjugate transpose (Hermitian) of a vector
$\lvert\lvert\cdot\rvert\rvert$	Euclidean norm
$\lvert\lvert\cdot\rvert\rvert_\infty$	maximum norm (maximum absolute value)
$\mathbb{C}^N$	N-dimensional complex Euclidean space
$\mathbb{R}^N$	N-dimensional real Euclidean space
$\mathbb{R}_+^N$	N-dimensional vectors whose elements are nonnegative.
B_N	all vectors in $\mathbb{C}^N$ with unit norm (i.e., unit N-ball)
$\otimes$	Kronecker product
$\odot$	Hadamard (entrywise) product
$\preceq$	element-wise inequality
$\circ$	function composition
$\mathcal{R}(\cdot)$	range of an operator
$\mathcal{N}(\cdot)$	kernel (null space) of an operator
$\mathbf{I}_N$	$N \times N$ identity matrix

Table 1: Mathematical Notation

We adopt the convention for complex partials and gradients described in [1, p. 105]. That is, let $x, y \in \mathbb{R}$ and $z \in \mathbb{C}$ be such that $z = x + jy$. Then the partials of $G : \mathbb{C} \rightarrow \mathbb{R}$ with respect to z and z^* are defined as

$$\frac{\partial G}{\partial z} := \frac{1}{2}\left\{\frac{\partial G}{\partial x} - j\frac{\partial G}{\partial y}\right\} \tag{4}$$

$$\frac{\partial G}{\partial z^*} := \frac{1}{2}\left\{\frac{\partial G}{\partial x} + j\frac{\partial G}{\partial y}\right\}. \tag{5}$$

Note that z and z^* are treated as independent variables of G. If $\mathbf{z} = [z_1 \ \ldots \ z_N]^T \in \mathbb{C}^N$, then the gradients of $G : \mathbb{C}^N \rightarrow \mathbb{R}$ with respect to $\mathbf{z}$ and $\mathbf{z}^H$ are defined as

$$\nabla_{\mathbf{z}} G := \left[\frac{\partial G}{\partial z_1} \ \frac{\partial G}{\partial z_2} \ \cdots \ \frac{\partial G}{\partial z_N}\right] \tag{6}$$

$$\nabla_{\mathbf{z}^H} G := \left[\frac{\partial G}{\partial z_1^*} \ \frac{\partial G}{\partial z_2^*} \ \cdots \ \frac{\partial G}{\partial z_N^*}\right]^T. \tag{7}$$

If a vector valued mapping $H : \mathbb{C}^N \rightarrow \mathbb{R}^N$ is defined by $H(\mathbf{x}) = [G_1(\mathbf{x}) \ldots G_N(\mathbf{x})]^T$, where $G_i : \mathbb{C}^N \rightarrow \mathbb{R}$, then the Jacobian of H with respect to the vector $\mathbf{x}$ is defined as

$$[J_{\mathbf{x}} H(\mathbf{x})]_{n,m} := [\nabla_{\mathbf{x}} G_n(\mathbf{x})]_m \tag{8}$$

For optimization packages that do not directly support complex valued design variables, we can exploit the conjugate symmetry between (6) and (7) to efficiently compute the gradient with respect to the real and imaginary parts of $\mathbf{z}$. That is,

$$\nabla_{\mathbf{z}_r} G = 2\mathrm{Re}\left\{\nabla_{\mathbf{z}} G\right\} \tag{9}$$

$$\nabla_{\mathbf{z}_i} G = -2\mathrm{Im}\left\{\nabla_{\mathbf{z}} G\right\}. \tag{10}$$

where $\mathbf{z}_r, \mathbf{z}_i \in \mathbb{R}^N$ and $\mathbf{z} = \mathbf{z}_r + j\mathbf{z}_i$ [2].

Chapter 1

Introduction

In this dissertation, we advance the theory of transmit adaptive (ATx) radar system design. Specifically, we present methods for generating transmit waveforms that a) maximize the detection probability in a known interference environment, and b) possesses certain characteristics required by the transmitter hardware and the receiver signal processing. The purpose of this chapter is to introduce transmit adaptivity (Sec. 1.1), place our work into the proper context (Sec. 1.2), and make explicit our contribution to the theory of ATx system design (Sec. 1.3).

1.1 Transmit Adaptivity in Radar

In this section, we define transmit adaptivity and discuss some of the potential benefits afforded by this technology (Sec. 1.1.1). We then discuss the many possible embodiments of transmit adaptivity (Sec. 1.1.2). This is followed by a more detailed discussion of the embodiment that is the subject of our research (Sec. 1.1.3).

1.1.1 Motivation

Radar systems transmit an electromagnetic signal into a volume of space containing both objects of interest (*targets*) and objects not of interest (*clutter*). The signal reflects from both kinds of objects, and these reflections are received by the radar along with signals produced by other sources (*interference*). The radar must then separate the target returns from the clutter and interference. There are currently effective techniques for adaptively suppressing clutter and interference at the receiver. Salient examples include: constant false alarm rate (CFAR) detectors, adaptive antenna beamforming techniques, and space-time adaptive processing (STAP). In theory, clutter and interference suppression could be enhanced if the transmit signal was tailored to the target/clutter/interference environment. However, despite potentially significant performance gains, this kind of transmit adaptivity has not yet become a mature technology. This is principally due to the insufficient computing power and limited RF waveform generation hardware that have heretofore rendered adaptive transmit techniques impractical to implement. However, the coupling of Moore's Law with recent advances in arbitrary waveform generation has dramatically improved the prospects of transmit adaptivity as a viable technology. As a result, there has been a resurgent interest in adaptive transmit radar systems. This is evidenced by a number of new conferences, research programs, and publications dedicated to transmit adaptivity [3–7].

So what does it mean exactly for a radar system to be transmit adaptive? A system is transmit adaptive if it is capable of altering its transmit waveform in response to knowledge about its environment. By *environment* we mean those elements that affect system performance, such as targets, clutter, and radio frequency interference (RFI). Knowledge of the environment could be acquired *a priori*, estimated online, or both. A notional ATx system is illustrated in Fig. 1.1. In general, a transmit waveform is designed using the environmental knowledge base. This designed waveform is then transmitted, and the transmitted signal reflects from both targets and clutter. The target and clutter reflections are collected at the receiver along with additive interference and receiver noise. Environmental

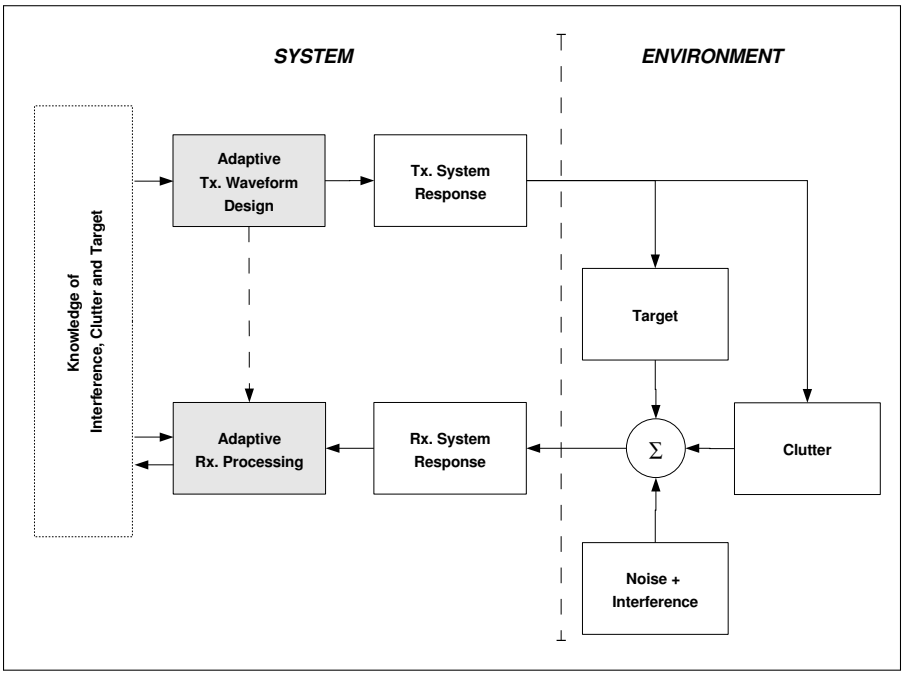


Figure 1.1: Functional block diagram of a transmit adaptive (ATx) system.

knowledge and knowledge of the transmit waveform are then used to adaptively process the received signal. The results of the signal processing stage might then be used to update the environmental knowledge base.

There are three main reasons to investigate transmit adaptivity in radar systems: 1) performance improvement, 2) resource management, and 3) novel missions. The first reason has to do with improving performance given constraints on the transmit waveform (e.g., peak power, nice autocorrelation), the remaining two reasons have to do with maintaining a minimum level of performance while either minimizing resources or performing multiple functions. For example, a transmit adaptive system might be able to compute a transmit waveform that maximizes the probability of detection in a given RFI environment (Reason 1). Such a system might also be able to maintain the probability of detection at a desired level while using a minimal amount of transmit power (Reason 2). Ideally, an ATx system would also be capable of performing two missions simultaneously, say spotlight synthetic aperture radar (SAR) and digital communications, while maintaining a minimum level of performance for each function (Reason 3). Of course, these are just notional examples.

However, they illustrate the gains that might accompany ATx technology. The research in this dissertation is concerned with Reason 1. That is, we present methods for finding waveforms to maximize the probability of detection subject to constraints on the transmit waveform.

1.1.2 Types of Transmit Adaptivity

While ATx has not yet developed into a mature technology, one should not conclude that the ATx research corpus is limited. In fact, ATx has been investigated for many applications (i.e., detection, estimation, tracking, imaging, classification). Furthermore, these investigations often have differing requirements regarding the environmental knowledge to be used and the degree of transmitter reconfigurability to be exploited. In order to place our work into the proper context, we find it useful to first define categories of ATx technology based upon the degree of reconfigurability required of the transmitter. We can then use this taxonomy to focus our discussion.

Figure 1.2 depicts a taxonomy of ATx methodologies based upon the degree of reconfigurability required of the transmitter. This taxonomy consists of two broad methodological families: *selection* and *design*. The selection family includes all methodologies that require waveforms, or waveform parameters, to be computed *before* operation. *Waveform selection* denotes those methodologies that adaptively choose a pre-designed waveform based upon current operating conditions. In a similar way, *parameter selection* denotes those methodologies that assume a canonical transmit waveform structure (e.g., pulsed LFM), but have the ability to adaptively select waveform parameters (e.g., pulse repetition frequency, carrier frequency, pulse width, chirp rate) from a pre-determined set. Methodologies in the design family, on the other hand, compute one or more waveform parameters *online*, that is, during operation. Methodologies belonging to the *parameter design* subfamily are an extension of parameter selection. These methodologies assume a particular waveform structure, but they do not select parameters from a pre-defined set. Instead, the values of

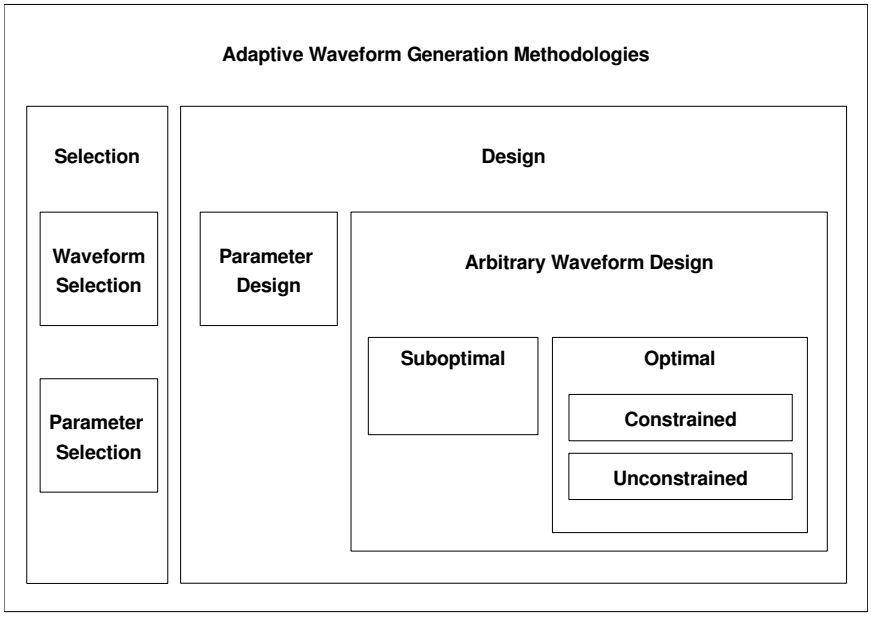


Figure 1.2: A taxonomy of ATx methodologies based on the degrees of freedom (DOF) required by the transmitter. The DOFs available to each methodology increases as we move from the top left corner to the lower right corner. An increase in the number of DOFs corresponds to an increase in potential performance improvement.

the parameters are calculated based upon information about the operating environment. Methodologies belonging to the *arbitrary waveform design* subfamily assume almost no structure for the transmit waveform. That is, each discrete-time sample is considered to be a degree of freedom (DOF) available to the design algorithm. Within the arbitrary waveform design subfamily belong *optimal* and *suboptimal* methodologies. Optimal methodologies directly maximize the radar performance metric of interest, whereas suboptimal methodologies apply a more heuristic approach to improving radar performance. Within the optimal methodologies we find *constrained* methodologies, which afford optimal performance (however it is defined) while explicitly taking into account constraints imposed by the system's transmit hardware and receiver signal processing. While an unconstrained algorithm is desirable, it is very rare that a problem formulation permits this.

It is not yet clear whether waveform selection or waveform design will be the more profitable approach to transmitter adaptivity. A shortcoming of the selection family is that

libraries of waveforms or waveform parameters must be designed before operation, and algorithms must search the library for the best performer. This can be computationally expensive. However, waveform library size can be quite small for some applications. For example, Moran *et al.* [8] considered a problem in which the transmitter is capable of transmitting only chirps of a given duration and bandwidth. Their work showed that the optimal library for their problem consists of just two waveforms – one with the greatest chirp rate, and one with the least. It was shown that waveforms with a chirp rate in between these endpoints are unnecessary. On the other hand, when the transmitter has a high degree of agility, as is assumed by arbitrary waveform design, then one expects waveform selection to under utilize the available DOFs compared to design methodologies. This could result in suboptimal performance. Furthermore, waveform design is more dynamic than waveform selection, and may therefore be able to make better use of dynamically changing environmental information.

The research presented in this dissertation belongs to the class of constrained optimal waveform design. That is, we seek the best possible solutions while considering all constraints placed on the waveform by the radar system. The benefit of this approach is that the constrained optimal waveform design is guaranteed to improve performance at least as well as any other method. Therefore, even if the solution method is inefficient, we have at least provided a tight upper bound on the performance gains that can be achieved. Optimal solutions are generally computationally expensive. As such, we also consider suboptimal designs that are computationally more attractive.

1.1.3 Arbitrary Waveform Design

By examining Fig. 1.1, we see that the arbitrary waveform design problem is comprised of two parts: *adaptation* and *optimization*. Furthermore, we recognize that adaptation can be decomposed into two principle activities, which are summarized by

$$\text{Adaptivity} = \text{learning} + \text{reconfiguring}.$$

That is, an ATx system must be capable of learning about its environment, and then reconfiguring its transmitter accordingly. The capability of such a system is limited by the degree to which it can perform either of these tasks. We have already addressed the idea of reconfiguration by choosing arbitrary waveform design as our application. However, to understand what the system must learn, and then address how well it can do so, we must first recognize that:

$$\text{Optimization} = \text{objective} + \text{constraints}.$$

An optimal waveform must maximize system performance while satisfying a number of system constraints. We call the measure of system performance the objective function. The objective function we use throughout this research is the signal-to-interference-plus-noise ration (SINR). However, other metrics can be found in the literature. (See Sec. 1.2.) Examples of system constraints that might be imposed on the waveform are 1) constant modulus, which is required for efficient transmission, and 2) a thumbtack ambiguity function, which is often required by the receiver signal processing scheme. In order to develop a viable ATx technology, one must be able to formulate system performance in terms of an objective function, and one must recognize all system constraints. Only then can the environmental parameters that must be inferred/estimated be identified (e.g., interference covariance matrix).

Our research is concerned with the optimization portion of the arbitrary waveform design problem, i.e., the *waveform optimization problem*. We assume that the adaptive problem (i.e., learning and reconfiguring) has been solved. This is not a particularly strong assumption because 1) arbitrary waveform generation is a mature technology, and 2) our work assumes that only the RFI power spectral density must be estimated, and this is a well understood problem in signal processing. We therefore focus on formulating the proper objective function and constraints, and we seek efficient solutions to the resulting optimization problem.

1.2 Literature Review

In this section, we review the ATx research literature. Our discussion begins with a justification of the scope of our review (Sec. 1.2.1). This is followed by a review of only the most relevant results (Sec. 1.2.2). A more detailed literature review is then presented in (Sec. 1.2.3). Note that the full review in Sec. 1.2.3 is chiefly for researchers of waveform optimization, and it may therefore be skipped by the more casual reader without losing the context of this work.

1.2.1 Scope

In the previous section, we restricted our discussion of adaptive transmit methodologies to the waveform optimization problem. However, the literature on waveform optimization addresses problems of target detection, parameter estimation, tracking, synthetic aperture radar (SAR) imaging, and target classification. It is not practical to discuss all of this literature at length. Since the proposed research is concerned with the detection of targets, we will restrict our review of the literature to only those works addressing the detection and classification problems. The classification literature is considered because the detection problem can be formulated as a special case of the classification problem. We will not discuss work concerning waveform optimization for target state estimation or tracking. This is because "improved detection performance also leads to improved tracking performance. This is due to the fact that, for a given probability of detection, a lower probability of false alarm implies less uncertainty in the origin of the measurement which leads to lower tracking error." [9].

In should also be noted that a great deal of waveform optimization research has occurred in the communications and controls literature. In fact, some problems in communications can be considered as duals of some radar waveform optimization problems, and early radar waveform optimization research borrowed techniques from the controls literature. A com-

plete review of the communications and controls literature is impractical. Instead, we shall discuss only those works that directly influence or benefit radar waveform design.

1.2.2 Most Relevant Results

Waveform optimization for clutter and interference suppression has been considered since at least 1965, when Van Trees remarked that "the most effective way (within the limitations of our model) to combat reverberation is through proper signal design." [10]. The earliest work focused on finding signal/filter pairs to maximize the signal-to-clutter-plus-interference ratio (SCIR) at the time of target detection [11–25]. These algorithms, which we refer to as *eigen-iterative* algorithms, iteratively solve a system of Fredholm integral equations by first finding the optimal filter for a given signal, and then finding the optimal signal for that filter, and then finding the optimal filter for that signal, and so on. In general, the only constraints placed on the signal and filter by the eigen-iterative algorithm are those of finite energy and duration. Amplitude and phase modulation limitations [16], implementation errors [19], constant modulus [23], and quantization effects [24, 25] were also addressed using extensions of the eigen-iterative approach. However, these individual constraints were never considered together in a single problem formulation.

The optimality of the eigen-iterative approach was never proven, but the point became somewhat moot when Sibul and Titlebaum [26] showed that simultaneous signal/filter design is unnecessary under the assumed signal model. They showed that, "the performance of the optimal receiver in detecting target echoes from a point target in signal-dependent interference depends only upon the channel power spectral density (PSD), the ambient noise PSD, the Doppler frequency shift induced by the point target's motion, and the power spectrum of transmit signal. When each of these factors is specified, the solution to the design of the optimal receiver is determined. Of all these factors, the only one that is under the control of an active sonar/radar system designer is the transmit signal." [27]. Under these conditions, the optimal receiver is determined by the Neyman-Pearson criteria. Taking

the Neyman-Pearson approach, Kay [28] extended work by Kooij [29], and used variational analysis to find the optimal transmit PSD under finite signal energy and bandwidth constraints. This approach has also been extended by Romero *et al.* [30] to consider non-moving stochastic extended targets. Romero *et al.* also provided similar solutions when mutual information is used as the performance metric [30]. Like Bell [31], they have found that the optimal waveform for detection generally differs from the optimal waveform for estimation/classification. Further, the optimal waveform for detection is a function of SINR.

The main drawback of [28] and [30] is that neither approach addresses the constraints placed on the waveform modulus and ambiguity function that commonly arise in practice. One approach to designing radar waveforms that satisfy these constraints is to formulate the optimal waveform and then minimize the the cost function subject to a set of constraints that *indirectly* constrain certain radar related properties of the waveform, such as the waveform modulus or autocorrelation sequence. This was done by Bergin *et al.* [32] by projecting a template (or desired) waveforms into the low noise subspace of the interference-plus-noise covariance matrix. Expanding on this work, Li *et al.* [33] devised an algorithm to calculate the optimal waveform given a maximum Euclidean distance between the solution and the template waveform. This *similarity constraint* has also been used by De Maio *et al.* [34–36].

The algorithms used to solve the constrained optimization problems considered in this dissertation are computationally expensive. As such, there is a perceived need for faster approximate solutions. Lindenfeld [37] used penalty functions to independently design the signal and receiver to enforce energy and autocorrelation constraints for the purpose of sparse frequency transmission. This approach could be used for interference suppression as well. Patton and Rigling [38] applied non-quadratic regularization in an attempt to constrain the waveform autocorrelation sequence.

1.2.3 Detailed Literature Review

Perhaps the most useful way to categorize the waveform optimization for target detection literature is according to the type of target model employed. The earliest work assumed a slowly varying point target with a narrowband approximation for Doppler, whereas a great deal of the later work would consider extended targets modeled as linear time invariant (LTI) systems. As we shall see, the earlier (point target) work was concerned primarily with detection performance, whereas the later LTI work was concerned with target identification. The research discussed in this dissertation is concerned with the detection of slowly varying point targets. However, we include a discussion on the extended target literature for completeness.

Point Target Models

Detection performance improvement via waveform optimization has been an ongoing topic of research since at least 1965 when Harry Van Trees remarked

> The most effective way (within the limitations of our model) to combat reverberation is through proper signal design. In fact, proper signal design is more important than optimum receiver design... An ideal system should have provision for continually measuring the scattering function and adapting the transmitted signal shape and receiver to the current environment. [10]

The design of VanTrees' ideal system can be seen as the subject of a flurry of work in the mid nineteen-sixties and early nineteen-seventies.[1] DeLong and Hofstetter [11, 12], Rummler [13], Ares [14], Thompson [15], Spafford [16], and Kincaid [17, 18] independently derived iterative algorithms for finding complex weights (amplitude and phase) for a coherent train of uniformly spaced pulses and the associated receive filter (i.e., a signal/-filter pair) that maximize SCIR at the time of target detection subject to a constraint on

[1] Note that waveform optimization for tracking performance improvement also began about this time [39]. However, we will not discuss this literature.

signal and filter energy. Spafford also considered signal amplitude and phase constraints. All of these algorithms iteratively solve a system of Fredholm integral equations by first finding the optimal filter for a given signal, then finding the optimal signal for that filter, and then finding the optimal filter for that signal, and so on.

In 1969, DeLong and Hofstetter [40, 41] derived a gradient method to design clutter resistant waveforms subject to limits on dynamic range. However, they later returned to the eigen-iterative approach [19] in order to address the design of radar signals/filter pairs subject to average energy and implementation error constraints. In the first textbook to describe adaptive radar waveform design, VanTrees [42, pp. 258] derived a coupled system of differential equations that describes the optimal signal for the detection of a slowly fluctuating point target in colored bandpass noise subject to energy and bandwidth constraints. Equations for an amplitude constraint are also discussed. However, solutions to these systems of equations are not provided. Zelio [20] compared the eigen-iterative signal/filter pair design of [11], and found that there were significant gains to be had over a traditional Tschebyscheff weighting. Lee [21, 22] developed an eigen-iterative algorithm that accounts for the time-varying multipath and dispersive ocean medium and interference environment with energy constraints on both the signal and filter. The resulting algorithm was very similar to [11]. Mesiya and Mclane [23] later applied [11] in conjunction with a least-squares approach to design signal/filter pairs, subject to a constant modulus constraint on the the signal, for the purpose of either maximizing signal-to-interference ratio or minimizing average sidelobe levels. Cohen [24, 25] considered Rummler's design procedure with the added restriction that the weights must be real and belong to a given finite set. The resulting nonlinear integer programming problem was solved efficiently using a branch and bound method.

While the eigen-iterative algorithms discussed above do converge, and they appear to provide performance improvement in simulation, they have not been proved to converge to a local optimum. Furthermore, Sibul and Titlebaum would show in 1981 that such iter-

ative algorithms are unnecessary when approached from a probability theoretic viewpoint instead of the more *ad hoc* SCIR viewpoint. To quote Sibul and Titlebaum [26]:

> The optimum (maximum-likelihood) receiver for detection of point targets in general, Gaussian interference is uniquely determined by the transmitted signal, white noise intensity, clutter (reverberation), scattering function, and target Doppler shift. The only function which is fully under the system designers control is the transmitted signal. Hence, the maximization of the detection index, or equivalently, maximization of the probability of detection for fixed false alarm probability, becomes a signal optimization problem. This is a difficult nonlinear optimization problem which requires a numerical solution. It is, however, a simpler solution than simultaneous optimization of two functions – the transmitted signal and receiver impulse response.

It should be noted that Balakrishnan [43] made a similar observation in 1968. However, his result was not the only important early result that would go unnoticed. In 1968, Kooij [29] reported an analytical solution to the problem of transmit signal optimization in the presence of signal-dependent clutter plus white ambient noise. He employed a calculus of variations approach to locally optimize the transmit signal power spectral density (PSD) subject to an energy constraint, but he did not attempt to show that his analytic solution was indeed a global maximum. This work by Kooij seems to have gone unnoticed until 2002 when Kay [27, 28] extended Kooij's work to provide an analytical solution to the problem of optimizing the transmit signal power spectrum to maximize the performance of an optimal receiver detecting echos from a single nonmoving Gaussian point target in the presence of signal-dependent clutter and colored ambient noise. Kay's work also addressed LTI target models. This will be discussed in the next section.

A unique early work (1969) was that of Holtzman [44] who proposed utilizing a weighted ambiguity surface for the performance criteria. Noting that the volume of the ambiguity function of a pulse train can be reduced if the pulses are non-uniformly modulated, Wong

and Chung [45] took an approach similar to Holtzman [44] and considered the minimization of the weighted distance of a multi-pulse ambiguity function from an ideal thumbtack ambiguity function – that is, they considered how the volume of the AF should be distributed. Genetic algorithms were then used to find optimum phase and frequency modulation for each pulse using weighted ambiguity volume [46] and average sidelobe volume [47] cost functions. It should be noted that this work by Wong and Chung did not allow arbitrary modulations, but only optimized certain *parameters* of the transmit waveform. Similarly, Bonneau and Wicks [48] sought the optimal chirp bandwidth for decorrelation with interference. In general, arbitrary waveform design should be more powerful, but these parameter design approaches may be more computationally tractable.

Another unique, relatively early work (1983), was that of Vastola [49], who considered the problem of uncertainty in clutter channel estimation. He devised a maximin approach to maximize the worst case performance, and provided a closed form solution for a particular clutter model.

In 2007, Sira *et al.* considered the adaptive design of radar waveforms for improved detection of small targets in heavy sea clutter [9, 50]. The authors note that previous waveform optimization work assumed the clutter returns were independent, and identically Gaussian distributed. However, the Gaussian model fails to predict the increased observance of higher amplitudes (spikes) for systems with spatial resolution fine enough to resolve the structure of the sea surface. This prompted the authors to use a compound-Gaussian (CG) model for sea clutter, which has (according to the authors) gained wide acceptance [51, 52] and has been tested both theoretically [53] as well as empirically [54]. The authors used clutter estimates to determine the autocorrelation lags at which the clutter was strong, and then used gradient methods to synthesize phase only signals that had low ACS sidelobes at those lags. This work was eventually extended [9] to the development of a particle filter based tracker that uses the measurements obtained by the adaptive waveform. A simulation study based on parameters derived from real sea clutter measurements demonstrated

that the approach provided a significant reduction in the tracking root mean square error (RMSE) when compared to a non-adaptive system.

The approach taken by Sira *et al.* [50] is "motivated by the fact that, in a radar, the signal obtained after matched filtering at the [matched] receiver is a convolution of the ambiguity function of the transmitted signal with the radar scene [55], which smears energy from one range-Doppler cell to another. Therefore, the incorporation of information about the clutter into the design of a waveform whose ambiguity function minimizes this smearing in the region of interest, can improve the SCR and detection performance." Thus, Sira *et al.* took a suboptimal approach to waveform design by finding signals with a property that is expected to characterize truly (i.e., provably) optimal waveforms. A similar kind of approach was taken by Holtzman [44] and Wong & Chung [45–47].

Linear Time-Invariant Target Models

In 1986, a monograph by Gjessing [56] entitled "Target adaptive matched illumination" was published. This book discussed the idea of matching the transmit waveform to the particular target being illuminated by the radar. As would be expected, A great deal of *a priori* information about the target is needed in order to apply matched illumination techniques. As such, matched illumination techniques are more often proposed for target characterization/identification than for detection.[2] There is extensive literature on the topic of matched illumination for target identification that began before Gjessing. Attempts at synthesizing signals for the purpose of target discrimination have been made since at least 1981 with the introduction of Kennaugh's K-pulse [58], which was soon followed by the E-pulse and the S-pulse [59]. These methods are based upon the singularity expansion method (SEM) in electromagnetics, which was introduced in 1971. SEM methods are interesting because they offer an *aspect-independent* means of representing a target's transient impulse response. However, these methods do have many practical limitations. The reader is

[2] In a review of Gjessing's book, Pell [57] states that the book might be more appropriately titled "target characterization with matched illumination applications."

referred to [59] for a good review of SEM based methods of signal synthesis.

In 1991, E.T. Jaynes [60] applied probability theory to the problem of discriminating between two targets with known deterministic impulse responses. He showed that "the transmitted pulse that is optimal for purposes of target discrimination will then have its spectrum concentrated near the frequency where [the absolute difference in target impulse spectra] reaches its absolute maximum." This idea was later expanded upon by Riggs *et al.* [61] who compared this theory with the SEM techniques discussed in [59].

The nineteen-nineties saw the first adaptive waveform design patents, which continue to be filed to this day [62–73]. Though not a target matched illumination scheme, Schreiber [62] patented an "adaptive waveform radar" that transmits only in the interference-free portion of the operating bandwidth. Target matched illumination for detection via SNR maximization was patented by Grieve and Guerci in 1992 [63–65].

The following year Bell [31] published a seminal work that addressed the adaptive waveform design algorithm using the information theoretic notion of mutual information. Bell considered the optimal waveforms for two problems involving the sensing of distributed Gaussian targets in additive white Gaussian noise (AWGN). The first problem was the design of the optimal detection waveform. The second problem was the design of an optimal estimation waveform. Interestingly, it was shown that the optimal detection waveform and the optimal estimation waveform (for the same distributed target) are quite different. The optimal detection waveform results from maximizing the SNR of the received signal whereas the optimal estimation waveform involves maximizing the mutual information between the transmit waveform and the target response. Bell motivated his use of mutual information as follows

> The mutual information $I(X;Y)$ between two random vectors X and Y tells us the quantity of information observation of Y provides about X ; that is, $I(X;Y)$ is the amount of information that the measurement Y provides about the object parameter vector X. The greater this mutual information is, the

> greater the quantity of information describing the object we obtain from our measurement and the greater the reduction in the *a priori* uncertainty as a result of this measurement. Intuitively, we might expect that the greater the mutual information between a measurement and the quantity being measured, the more accurately we can classify or estimate the parameters characterizing the entity we are trying to measure. Berger alludes to this idea when he makes the statement: Rate distortion theory provides knowledge about how the frequency of faulty categorization will vary with the number and quality of observations [Berger, p. 91], since the rate distortion function relates the average distortion or error to the minimum mutual information required to achieve that error.

Bell's paper can be seen as the foundational work for all information theoretic waveform design algorithms that were to follow.

It should be noted that the problem of target matched illumination for target identification is a dual of the channel matched signal design problem in communications. (See [74, 75] for an example.) The duality between the communications and radar problems is most apparent in Sowelam and Tewfik [76], who in 1994 proposed a method of designing a set of waveforms for the identification and subsequent imaging of targets with known responses. The same basic idea was patented by Guerci [66] in 1995. Barrett [68] patented target matched illumination with considerations to the effects of the transmission medium. Gjessing [77] discussed target matched illumination for naval applications. Pillai *et al.* [78, 79] [80, ch. 8], apparently unaware of the previous work on this topic [11–13, 15–18, 22], developed an eigen-iterative algorithm for optimal signal/filter pairs for the case in which the target and clutter can both be modeled as linear time invariant random processes.

It was soon realized that the approach in [79] could be re-formulated for the target identification problem by relating SCIR to the Mahalanobis distance [81–83]. Garren *et al.* applied [79] to the target identification problem using synthetic (X-patch) target data.

Specifically, they assumed perfect knowledge of the aspect-dependent transfer functions of two targets (a T72 and an M1 battle tank), with [84, 85] and without [86] some uncertainty in azimuth aspect. Their results showed that performance quickly degrades when aspect is unknown. The approach in [79] was extended to the multichannel scenario by Pillai *et al.* [87]. Again, Garren *et al.* applied the result using synthetic signature data [88]. Their results seemed to indicate that the multichannel (in this case, polarization) extensions to the algorithm does not aid much in the detection and identification of targets. The multi-target identification problem was later discussed by Pillai *et al.* in [89], and eventually patented [70, 71] in 2006. The multichannel approach was later re-cast for the MIMO waveform design problem by Pillai *et al.* in 2003 [90].

Despite all of the extensions to [79], the proposed algorithm is neither guaranteed to converge nor produce a locally optimal signal. Further, it is not clear that this algorithm performs better than the earlier eigen-iterative algorithms, which were, incidentally, being studied by Haykin *et al.* [91, 92] at about the same time. However, the efficacy of this approach is a moot point because these iterative approaches were shown to be unnecessary by Sibul and Titlebaum [26]. As Kay [27, 28] remarked:

> The optimal receiver is usually defined as one which maximizes the probability of detecting a target in the presence of interference (i.e., reverberation plus ambient noise) for a specified probability of false alarm. The solution for the design of such an optimal receiver, given a statistical characterization of the interference, is available in a number of standard texts [42, 93]. It can be shown that the detection performance of the optimal receiver is a monotonic function of the signal to interference ratio (SIR). When the interference is white in frequency, the SIR depends only on the total energy of the transmit signal. However, when the interference is not white, and furthermore, signal-dependent, the design of the transmit signal can have a significant impact upon the detection performance of the optimal receiver.

> It was pointed out two decades ago [26] that the performance of the optimal receiver in detecting target echos from a point target in signal-dependent interference depends only upon the channel power spectral density (PSD), the ambient noise PSD, the Doppler frequency shift induced by the point target's motion, and the power spectrum of the transmit signal. When each of these factors is specified, the solution to the design of the optimal receiver is determined. Of all these factors, the only one that is under the control of an active sonar/radar system designer is the transmit signal.

In 2002, Kay [27, 28] extended Kooij's 1968 work [29] to provide an analytical solution to the problem of optimizing the transmit signal power spectrum to maximize the performance of an optimal receiver detecting echos from a single nonmoving Gaussian point target in the presence of signal-dependent clutter and colored ambient noise. It should be noted that the restriction to Gaussian point targets was made for pedagogical purposes only, with the extension to extended (LTI) targets like those previously considered being a straight forward extension. Like previous work, Kay's paper assumes the target and clutter act as stochastic LTI systems. Kay remarked about the limitations of these LTI models [28]:

> The scattering model assumes that the signal-dependent noise is the output of a random linear time invariant (LTI) filter, whose impulse response can be assumed to be a realization of a wide sense stationary (WSS) random process. The same model has been used before in [31] and more recently in [87]. It should be noted that this model does not allow for spectral spreading, as would be inherent in a moving platform and/or intrinsic clutter motion situation. Hence, it differs from the standard one usually assumed [42].

However, subject to the limitations above, the advantage of such an LTI model is that 1) an analytical solution for the optimal waveform is obtained, and 2) new insights into the signal design problem are evident.

In 2004, Tan and Zheng [94] considered a somewhat different version of the aforementioned problem, i.e., that of minimizing the estimation error of the target profile. They used the Wiener filter to reconstruct the target profile in the MMSE sense, and the transmit signal was optimized (subject to pulse energy constraint) to further minimize the mean square error achieved by the Wiener filter. Like previous work, their model assumes a priori knowledge of the target and noise in the form of their power spectral densities. They extend this work to multiple pulse transmission and deconvolution in that, after each pulse, the *a posteriori* PSD is calculated and used as the *a priori* PSD of the next pulse.

In 2000, Sowelam and Tewfik [95] developed and information theoretic approach to target identification based upon synthesis of a sequence of probing signals to maximize classification performance. Their approach can essentially be treated in the context of experimental design, which is an organized method for extracting as much information as possible from a limited number of observations. The information theoretic sequential hypothesis testing approach was also taken by Goodman *et al.* [96–98]. This approach was extended by Butler and Goodman to bistatic MIMO scenarios [99, 100]. Romero, Bae, and Goodman also compared SNR and MI-based waveform design in [30]. Like Bell, they found that the SNR and MI-based waveforms are generally different for given target and clutter impulse responses.

Constraint Satisfaction

The optimal PSD solutions of [28] and [30] are valid under finite signal energy and bandwidth constraints. However, the resulting time domain representations of such waveforms are not practical. First, the finite bandwidth constraint forces the signal to be of infinite duration. Second, the energy constraint is insufficient to guarantee that the signal will satisfy common envelope requirements. Even when stringent envelope requirements such as constant modulus are not imposed, the peak power of the waveform must still be constrained because scaling processes that takes place in the system can cause the otherwise optimized

waveform to be detrimental to system performance. (See Ch. 2). Consistent with other problem formulations (e.g., [31, 79]), the solutions of [28] and [30] assume that only a single target is present in the scene, and that the target range and Doppler shift are known. When this is not the case, the signal/filter cross-ambiguity function (CAF) becomes an important driver of system performance. In general, the CAF is required to have a narrow main lobe so that closely spaced targets can be resolved, and low side lobes to reduce clutter returns and prevent weaker target returns from being masked by stronger ones [101–103]. (See Ch. 3).

One approach to designing radar waveforms that satisfy modulus constraints is to formulate the optimal waveform and then minimize the the cost function subject to a set of constraints that *indirectly* constrain certain radar related properties of the waveform, such as the waveform modulus or autocorrelation sequence. This was done by Bergin *et al.* [32] for the detection of known extended targets in colored interference. Their method was to project a template (or desired) waveforms into the low noise subspace of the interference-plus-noise covariance matrix. The reasoning behind this approach is that the resultant waveform will be similar to the desired waveform in the sense that the two waveforms are as close as possible in a Euclidean sense. Li *et al.* [33] later devised an algorithm that provides the optimal waveform given a maximum Euclidean distance between the solution and the template waveform. This *similarity constraint* has also been used by De Maio *et al.* [34–36]. (See Ch. 4.)

In general, solutions to the constrained optimization problems considered here are computationally expensive. That is, the time required to compute a solution may be longer than the duration over which the statistical assumptions are valid. As such, there is a perceived need for approximate solutions. Lindenfeld [37] used penalty functions to independently design the signal and receiver to enforce energy and autocorrelation constraints for the purpose of sparse frequency transmission. Note that his approach can be used for interference suppression as well. Patton and Rigling [38] applied non-quadratic regularization in an

attempt to constrain the autocorrelation sequence of a matched filtered signal.

1.3 Contributions

This dissertation advances the theory of adaptive transmit radar technology by addressing the satisfaction of modulus and ambiguity function constraints in the target detection problem. Modulus constraints are addressed in Ch. 2. In Sec. 2.1, we demonstrate that even when the transmitter does not require the waveform to have a constant envelope, the waveform should still be optimized subject to peak power constraints. Algorithms that ignore these constraints risk producing designs that can be detrimental to system performance. These results were originally published in [104]. In Sec. 2.3, we consider the problem of extending optimal designs such as [28] and [30] to handle duration and modulus constraints. We show that this is a problem of phase retrieval, which can be solved via dynamic programming (i.e., gradient-based search methods) or the method of alternating projections. We extend current phase-retrieval techniques, and show that the alternating projections approach can be an efficient alternative to dynamic programming. These results were originally reported in [105].

In Ch. 3, we consider the multiple target detection problem, which requires the signal/-filter cross-ambiguity function to be constrained. For simplicity, we explicitly deal only the signal/filter cross-correlation sequence (XCS). However, extending these results to the cross-ambiguity function is a conceptually simple exercise. We formulate the waveform optimization problem for a number of receiver architectures, and we show that, unlike in the single target detection problem, the Neyman-Pearson filter may not be the optimal receive architecture. These results were originally reported in [106].

When XCS constraints are involved, the algorithms for solving the waveform optimization problem can be computationally demanding. In Ch. 4 we examine several formulations that require fewer calculations. In Sec. 4.1, we consider six different formulations and find

that a phase-only waveform design can provide nearly optimal results with a much reduced computational complexity. We also show that the similarity constraint affords computationally efficient solutions, but at the cost of greatly reduced gains. These results were originally reported in [107]. In Sec. 4.2, we design an algorithm for indirectly constraining the waveform ACS in a phase-only design. This algorithm is capable of providing greater gains than the similarity constraint approach, but at a greatly reduced computational burden compared to the direct constraint approach in Ch. 3. These results were originally reported in [108].

Chapter 2

Modulus Constraints

The variational solution in [28] and the SWORD solution in [33] both produce globally optimal waveforms for their respective constraint sets. However, neither approach constrains the modulus (complex envelope) of the transmit waveform. This is problematic because radar systems commonly require constant modulus (i.e., constant amplitude) waveforms so as to efficiently utilize nonlinear power amplifiers. Furthermore, we show in Sec. 2.1 that the waveform modulus must be constrained even when linear amplifiers are used. This analytic result is illustrated by a numeric example in Sec. 2.2.

After establishing the inextricable nature of modulus constraints in the waveform optimization problem, we dedicate the remainder of the chapter to extending the state of the art in waveform design to accommodate modulus constraints. This amounts to finding a waveform that satisfies the time domain constraints and has the optimal power spectral density. In Sec. 2.3, we connect this problem to the problem of phase retrieval. We then develop solutions based on dynamic programming and alternating projections (Sec. 2.4-2.6). Numerical examples in Sec. 2.7 demonstrate that the alternating projections approach can provide accurate solutions at a lower computational cost than dynamic programming approaches. Concluding remarks for this chapter are given in Sec. 2.8.

2.1 Waveform Scaling

The waveform optimization algorithms presently found in the literature will necessarily be implemented on a digital computer, and the resulting waveforms will be converted to analog form by a digital-to-analog converter (DAC). This conversion process is depicted in Fig. 2.1. Let us assume that the amplifier is linear, and that the desired output voltage is $\pm A$. Working backwards through the chain, we see that if the amplifier's gain is G_2, then the input voltage to the amplifier must be $\pm A/G_2$. Similarly, if the linear DAC has gain G_1, then the DAC input voltage must be $\pm A/(G_1 G_2)$. If a waveform optimization algorithm does not explicitly constrain the waveform to have a maximum amplitude of $\pm A/(G_1 G_2)$, then the optimized waveform must be scaled before being fed to the DAC. If we let $\mathbf{s}$ denote the transmit waveform, then the scaled waveform is given by

$$\overline{\mathbf{s}} := \frac{A/(G_1 G_2)}{||\mathbf{s}||_\infty}\mathbf{s}. \tag{2.1}$$

In order to illustrate the effects of waveform scaling on the performance of a waveform optimization algorithm, consider the problem of detecting a known signal in the presence of colored Gaussian noise using a matched filter detector. As we will see in Ch. 3, the optimal waveform will maximize the post-filter signal to noise ratio, which is given by

$$\text{SINR}\left(\mathbf{s}\right) = \frac{\left|\mathbf{s}^H\mathbf{s}\right|^2}{\mathbf{s}^H\mathbf{K}\mathbf{s}}, \tag{2.2}$$

where $\mathbf{s} \in \mathbb{C}^N$ is the transmit signal and $\mathbf{K} \in \mathbb{C}^{N\times N}$ is the interference covariance matrix. Furthermore, the performance of a waveform optimization algorithm is typically measured by the gain it provides with respect to some baseline solution. Letting $\mathbf{s}_0$ denote the baseline solution, we can write the gain as

$$\text{GAIN}\left(\mathbf{s}\right) := \frac{\text{SINR}\left(\mathbf{s}\right)}{\text{SINR}\left(\mathbf{s}_0\right)}. \tag{2.3}$$

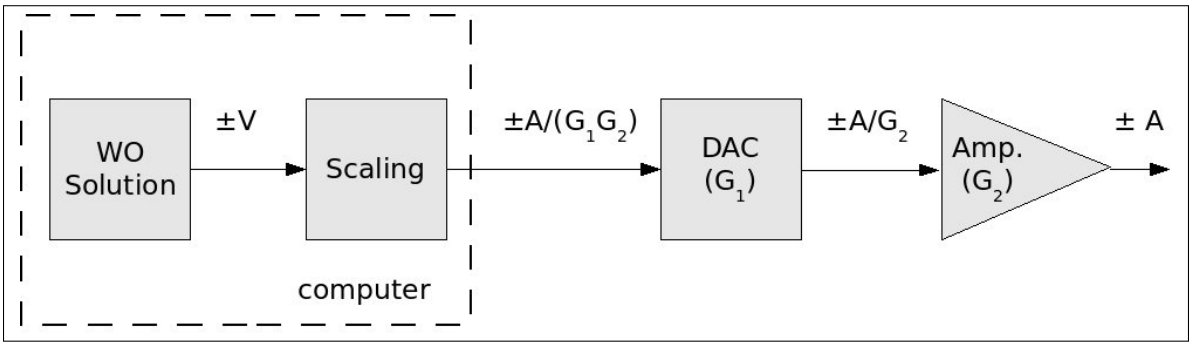


Figure 2.1: Functional block diagram of the waveform scaling process.

However, the true transmit waveform will be scaled, and the actual gain will be

$$\frac{\text{SINR}\left(\overline{\mathbf{s}}\right)}{\text{SINR}\left(\overline{\mathbf{s}}_0\right)} = \frac{||\mathbf{s}_0||_\infty^2}{||\mathbf{s}||_\infty^2}\text{GAIN}\left(\mathbf{s}\right). \tag{2.4}$$

From this equation we see that the optimized waveform will improve performance above the baseline solution (i.e., the true gain will be greater than unity) only if

$$||\mathbf{s}||_\infty \leq ||\mathbf{s}_0||_\infty \sqrt{\text{GAIN}\left(\mathbf{s}\right)}. \tag{2.5}$$

This relationship can be interpreted as stating that the waveform can put less actual power on the target if the gain provided by spectrum shaping is sufficiently large. In order to ensure the optimized waveform satisfies (2.5), the optimization algorithm should instead maximize the scaled gain. This is equivalent to maximizing

$$\overline{\text{SINR}}\left(\mathbf{s}\right) := \frac{1}{||\mathbf{s}||_\infty^2}\text{SINR}\left(\mathbf{s}\right). \tag{2.6}$$

Alternately, the algorithm could maximize the SINR in (2.2) subject to the constraint $||\mathbf{s}||_\infty \leq ||\mathbf{s}_0||_\infty$.

From the above discussion, we see that the scaling process imposes a maximum modulus (or peak power) constraint on the optimized waveform. Algorithms that do not in-

corporate this constraint risk performance degradation because the scaled waveform may put less than the expected energy on the target. If the waveform is not scaled, the output of the DAC will be a clipped version of the input signal, and system performance will be degraded because the transmit waveform will not be the same as the optimized waveform.

The deleterious effect waveform scaling can have on system performance demands it be accounted for by the waveform design algorithm. However, many arbitrary waveform design algorithms constrain only the energy of the transmit waveform. While an energy constraint is reasonable, it is insufficient. This is because the *amount* of energy is constrained while the *distribution* of that energy throughout the pulse is not. This can lead to designed waveforms with large modulus variations (i.e., large dynamic range), and scaling can be particularly pernicious to these kinds of waveforms. We should note that nonlinear RF power amplifiers are often used in radar systems, and these amplifiers require the input signal to be constant modulus [102, 109]. The constant modulus constraint is even more extreme than the peak power constraint discussed in this section, and it should be accounted for by the waveform optimization algorithm.

2.2 Example: Ignoring Modulus Constraints

In order to illustrate the problems that can arise when the maximum modulus constraint imposed by the scaling process is ignored, we simulate the scaling process applied to waveforms generated by four waveform optimization algorithms found in the literature. Specifically, we consider the three eigenvector-based solutions in [32] (denoted $M_1 - M_3$, respectively), and the SWORD algorithm in [33]. These algorithms were designed for the problem of detecting a known signal in the presence of wide-sense stationary colored Gaussian noise using a Neyman-Pearson whitening filter detector (see Ch. 3). Each algorithm

maximizes the post-filter SINR, which is given by

$$\text{SINR}\left(\mathbf{s}\right) = \mathbf{s}^H \mathbf{K}^{-1} \mathbf{s} \tag{2.7}$$

when the whitening filter is used. The differences between the algorithms have to do with the constraints to which this maximization is subjected. These constraints are not particularly relevant to our purposes here. So, we will not discuss the algorithms in detail. However, we do note that each algorithm requires the waveform to have unit energy, but no constraints are placed on the waveform modulus.

The performance of these methods when *not* subjected to scaling was simulated in [33] for an example interference process consisting of a white Gaussian random process (i.e., the noise) added to an autoregressive random process (i.e., the interference) whose transfer function is given by

$$H(z) = \frac{1}{(1 - 1.5z^{-1} + 0.7z^{-2})^4}. \tag{2.8}$$

We have reproduced those simulations here. The results are shown in Fig. 2.2, which depicts SINR improvement (gain) versus interference-to-noise ratio (INR) for the various waveform design algorithms. For this example, the baseline solution was a critically sampled complex-valued linear frequency modulated (LFM) waveform with unit energy. Notice that, as expected, all methods provide positive gain when not subjected to scaling.

Fig. 2.3 shows SINR improvement versus INR for the same scenario with the exception that the optimized waveforms were subjected to scaling. In this case, the maximum DAC input, which we will denote by M_{in}, was specified to be $M_{in} = 1/\sqrt{N}$.[1] Notice that the scaled waveforms actually provide a negative gain when subjected to scaling. That is, the non-optimized baseline solution provides better system performance. The reason for

[1]Note that this scaling forces the maximum modulus of the designed waveform to equal $1/\sqrt{N}$. This corresponds to unit energy for a constant modulus waveform. Also, for this simulation, $N = 301$.

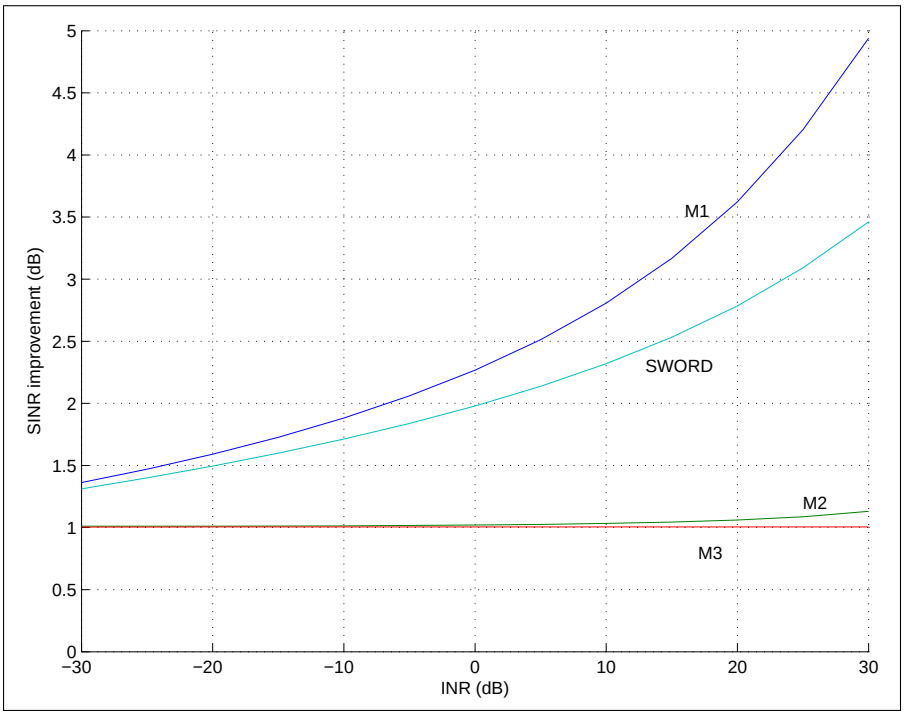


Figure 2.2: Gain vs. INR when waveforms were not subjected to scaling.

performance loss is clear when we examine Figs. 2.4 and 2.5, which depict the modulus (complex magnitude) of each waveform for the INR = 20 dB case. Fig. 2.4 shows the unit energy waveforms before scaling, and Fig. 2.5 shows the waveforms after scaling. The black dashed line represents M_{in}. For this example, the scaling process resulted in a transmit power reduction of approximately 3, 20, 3, and 5 dB for $M1$, $M2$, $M3$, and SWORD, respectively. These losses are significant compared to the gains resulting from the optimization process (see Fig. 2.2), and the resulting performance is commensurate with the relationship in (2.4).

2.3 Phase Retrieval for Radar Waveform Optimization

As we have already established, radar waveform optimization involves finding a transmit waveform that optimizes system performance and satisfies system constraints. And, as we have seen, there has been extensive work done on the design of waveforms for the

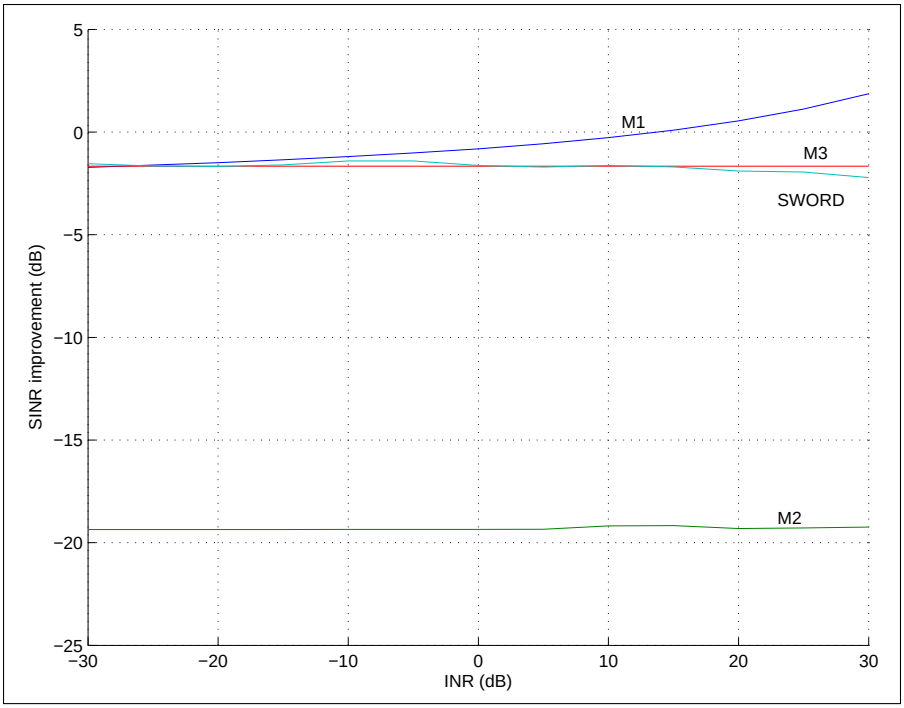


Figure 2.3: Gain vs. INR when waveforms were subjected to scaling.

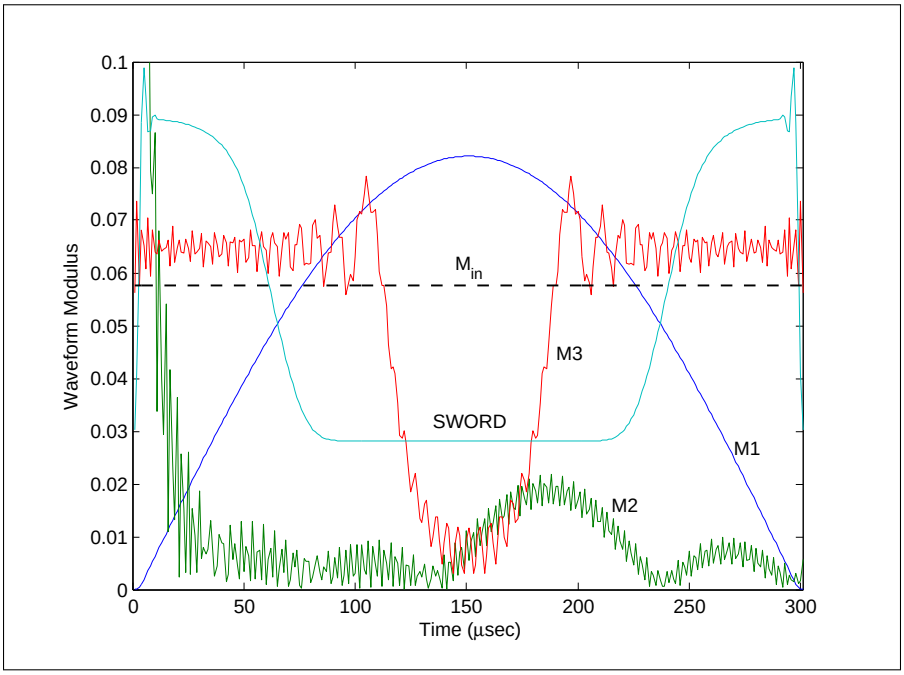


Figure 2.4: Moduli of optimized waveforms before scaling. Note that each waveform must be scaled so that its maximum modulus is no greater than the dashed line – that is, no greater than $M_{in} = 1/\sqrt{N}$.

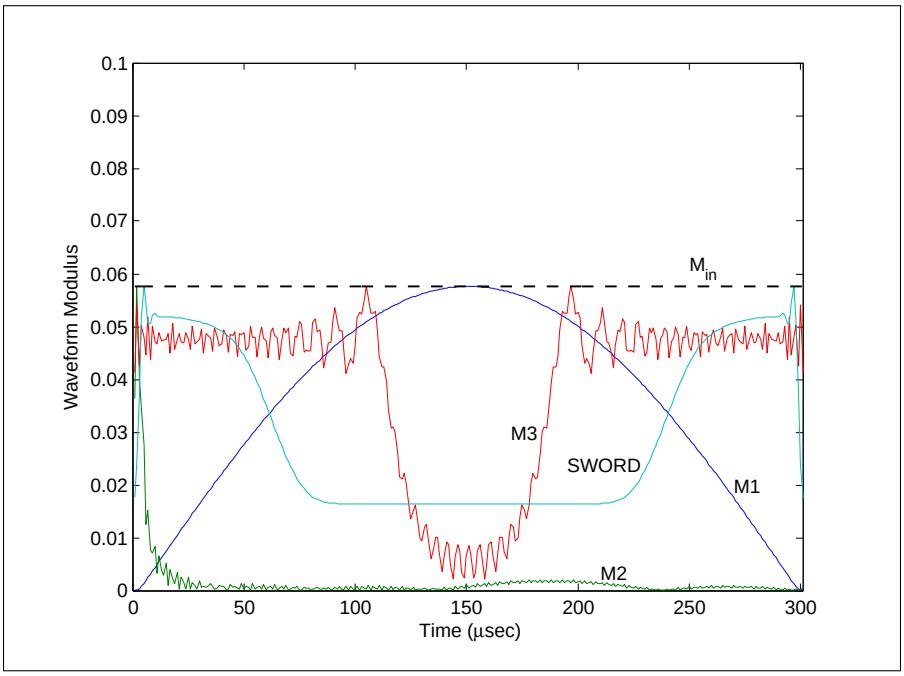


Figure 2.5: Moduli of optimized waveforms after scaling. Note that scaling resulted in a power reduction of apx. 3.1, 20.4, 2.7, and 4.7 dB for $M1$, $M2$, $M3$, and SWORD respectively.

minimization of mean squared error (MSE) in SAR image reconstruction [110], the maximization of signal-to-noise ratio for target detection [28, 79, 111], and the maximization of mutual information [31, 98, 112] or Mahalanobis distance [86, 88] for target classification. Furthermore, in each of the works cited above, it can be shown that when the only constraints placed on the transmit waveform are finite energy and bandwidth, the transmit waveform affects system performance only through its Fourier transform magnitude (FTM) function.[2] Convenient expressions for optimal FTMs do exist for such cases (e.g., [28]). However, the inverse Fourier transforms of these FTMs are not practical radar waveforms. This is because the finite bandwidth constraint forces the signal to be of infinite duration, and the energy constraint is insufficient to guarantee that the signal will satisfy practical envelope constraints that arise from the use of nonlinear amplifiers (i.e., constant modulus constraints), or from the scaling processes discussed in the previous section (i.e., peak power constraints). Some applications may even impose constraints (e.g., Doppler toler-

[2]That is, the magnitude of the Fourier transform.

ance) that cannot be addressed by the finite energy/bandwidth constraints alone. Therefore, it is desirable to find waveforms of the appropriate duration that satisfying the modulus constraints and minimize the MSE between the actual and optimal FTMs. These waveforms may be practical for some applications, and will in any case provide better performance bounds than the finite energy/bandwidth constrained waveforms.

The problem of finding waveforms of the appropriate duration that possess both the desired time domain envelope and the optimal FTM is a problem of phase retrieval. Phase retrieval problems have been studied extensively in many fields such as electron microscopy, wavefront sensing, astronomy, crystallography, signal processing, and image processing. (See references in the works cited below for examples.) In the general phase retrieval problem, one knows the modulus of the Fourier transform of an "object," as well as a set of *a priori* constraints on that object. The problem is then to find the Fourier phase function that produces an object satisfying those constraints. To put the problem in mathematical terms, let $\mathcal{F} : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ denote the Fourier transform from space $\mathcal{H}_1$ (known as the *object domain*) to space $\mathcal{H}_2$ (the *image domain*). Let $A \subseteq \mathcal{H}_1$ denote the object domain vectors that satisfy all object domain constraints (e.g., constant modulus), and let $\widetilde{B} \subseteq \mathcal{H}_2$ denote the image domain vectors that satisfy all image domain constraints (e.g., possess the optimal FTM). The phase retrieval problem becomes that of finding an $\mathbf{a} \in A$ such that $\mathcal{F}\mathbf{a} \in \widetilde{B}$. Note that a particular phase retrieval problem is completely characterized by the constraint sets A and $\widetilde{B}$. In practice, these sets are often *inconsistent*, meaning that there is no $\mathbf{a} \in A$ such that $\mathcal{F}\mathbf{a} \in \widetilde{B}$. In this case, one often seeks an object whose Fourier transform comes as close as possible to satisfying the image domain constraints.

In general, a phase retrieval problem will not admit a closed form solution, but it can usually be solved using an iterative transform method (ITM). Iterative transform methods repeatedly transform between the object and image domains, satisfying the constraints in one domain before returning to the other [113]. While this approach is perhaps physically intuitive, explaining its success over such a broad class of applications is still an open

problem [114–119]. One thing appears clear, however, the effectiveness of ITMs is the result of their alternating projections (AP) nature.

The phase retrieval problem with which we are concerned has arisen in other fields under the names *spectrum shaping* [120], *two-intensity measurement problem* [113], and *phase-only pulse shaping* [121]. In fact, the earliest ITM, called the Gerchberg-Saxton algorithm (GSA), was developed specifically for this kind of problem [122]. Subsequent development of ITMs, however, was motivated by issues of efficiency, solution quality, and algorithm stability in another phase retrieval problem known as the *single-intensity measurement problem* [113]. In a seminal work, Fienup [113] showed the GSA to be a special case of the error reduction algorithm (ERA). He then used the ERA framework to develop the basic input-output (BIO) algorithm for convergence time reduction, and the hybrid input-output (HIO) algorithm to reduce the occurrence of convergence stagnations [113]. Youla and Webb [123] recognized the role of alternating projections in iterative transform algorithms, but their analysis was performed only for consistent problems involving convex constraint sets. Levi and Stark [124] extended this approach to consider nonconvex constraints, and implicitly addressed inconsistent problems as well. The AP approach was later used by Elser [125] to develop the difference map algorithm, by Bauschke *et al.* [115] to develop the hybrid projection reflection (HPR) algorithm, and most recently by Luke [117] to develop the relaxed averaged alternating reflection (RAAR) algorithm. It should be noted that the spectrum shaping problem does not share the same stability issues as the single-intensity measurement problem, which motivated the development of these other algorithms [113]. Thus, these algorithms may provide little advantage over the GSA. Furthermore, the GSA is parameter free, whereas the other algorithms require parameters that cannot be determined *a priori* [117, 124]. For these reasons, the GSA is the most practical ITM for the spectrum shaping problem.

Inconsistent phase retrieval problems can also be formulated as constrained optimization problems, which can be solved using more broadly applicable techniques. For ex-

ample, Fienup [113] found that the performance of the ERA, BIO and HIO compared favorably to that of a constrained gradient descent (GD) approach for the single-intensity measurement problem, and he demonstrated some similarities between the GSA and GD for the spectrum shaping problem. Rundquist [121] compared the GSA to a genetic algorithm for a spectrum shaping problem, and found that the GSA generally provided a better solution while converging 10 to 100 times faster [121]. Hacker reports similar results with respect to a genetic algorithm and a combination of a downhill simplex method and simulated annealing [126].

Several researchers have recently applied the GSA to the radar waveform synthesis problem described above [127–129]. However, several important issues remain unaddressed. First, in typical discrete phase retrieval problems, the length of the time domain vector is equal to the number of FTM samples. This need not be so in waveform synthesis, but the case of unequal vector lengths has not been thoroughly studied, especially when the length of the time domain waveform is greater than the length of the FTM vector. Second, it is unclear how the GSA compares in terms of performance and efficiency to more standard nonlinear programming methods (e.g., gradient descent or quasi-Newton methods) when the time and frequency domain vectors are not of equal length. Both of these questions are addressed in the present work.

In this chapter, we adopt the alternating projections framework to explain the success of the GSA in previously studied cases, and to extend the algorithm for previously unstudied scenarios. We explain convergence behavior for the general case of unequal vector lengths by proving the so-called error reduction property, and we present numerical results demonstrating that the alternating projections approach can provide nearly identical performance, at a much lower computational cost, than competing nonlinear programming algorithms. Further, the AP approach is simpler to implement than competing nonlinear programming approaches, making it a convenient choice for many applications.

The remainder of this chapter is structured as follows. The alternating projections na-

ture of the error reduction algorithm, of which the GSA is a special case, is discussed in Sec. 2.4. This discussion motivates the algorithm development in Sec. 2.5. The discrete FTM waveform synthesis problem is defined first, and we derive an alternating projections equivalent of the GSA in Sec. 2.5.1. In Sec. 2.5.2, we derive an alternating projections algorithm that is not equivalent to the GSA. The similarities between the alternating projections algorithm and the gradient descent algorithm are discussed in Sec. 2.6. Section 2.6.3 contains a brief discussion of the problem in which the signal's power spectral density (PSD) is of greater importance than its FTM. Numerical examples are presented in Sec. 2.7, and Sec. 2.8 contains concluding remarks. Note that all proofs and derivations are relegated to the Appendix.

2.4 The Error Reduction Algorithm

In this section, we discuss the alternating projections nature of the ERA so that we might understand the convergence behavior of the GSA, and motivate the algorithm development that follows. To begin, consider the inconsistent problem in which we seek an $\mathbf{a} \in A \subseteq \mathcal{H}_1$ such that $\mathcal{F}\mathbf{a} \in \widetilde{B} \subseteq \mathcal{H}_2$, but no such $\mathbf{a}$ exists. In such cases, one usually seeks an $\mathbf{a} \in A$ whose Fourier transform $\mathcal{F}\mathbf{a} \in \mathcal{H}_2$ is as close as possible to the image domain constraint set $\widetilde{B} \subseteq \mathcal{H}_2$ [113, 124]. This is equivalent to minimizing the point-set distance $d(\mathcal{F}\mathbf{a}, \widetilde{B})$, which is referred to as the *error*. The ERA seeks to iteratively minimize the error through four steps: 1) Fourier transform the current object domain vector. 2) Make the *minimum changes* to the transformed vector so that it satisfies the image domain constraints. 3) Inverse Fourier transform the resulting vector. 4) Make the *minimum changes* to the inverse transformed vector so that it satisfies the object domain constraints. These four steps are repeated until convergence is achieved. The need for *minimal changes* in steps 2 and 4 implies that the ERA is only applicable if there exists projector mappings $P_A : \mathcal{H}_1 \rightarrow A$ and $P_{\widetilde{B}} : \mathcal{H}_2 \rightarrow \widetilde{B}$ such that $P_A(\mathbf{x}) \in \Pi_A(\mathbf{x})$ and $P_{\widetilde{B}}(\tilde{\mathbf{x}}) \in \Pi_{\widetilde{B}}(\tilde{\mathbf{x}})$ for all $\mathbf{x} \in \mathcal{H}_1$

and $\tilde{\mathbf{x}} \in \mathcal{H}_2$.That is, both constraint sets must have well defined projections from which a single-valued selection can be made. Using this notation, we can write the ERA as in Alg. 1. Here, algorithm lines 4-7 correspond directly to steps 1-4 above.

We now examine the ERA in terms of alternating projections to render a more succinct update expression, and to also illustrate the error reduction property. Let $B \subseteq \mathcal{H}_1$ be the set of object domain vectors whose Fourier transform satisfies the image domain constraint. Since $\mathcal{F}$ is unitary, we have $\widetilde{B} = \mathcal{F}B$ and

$$d(\mathbf{a}_k, B) = d(\mathcal{F}\mathbf{a}_k, \mathcal{F}B) = d(\mathcal{F}\mathbf{a}_k, \widetilde{B}) \tag{2.9}$$

Further, the mapping $P_B : \mathcal{H}_1 \rightarrow B$ defined by

$$P_B(\mathbf{x}) = \mathcal{F}^{-1} P_{\widetilde{B}}(\mathcal{F}\mathbf{x}) \tag{2.10}$$

satisfies the condition $P_B(\mathbf{x}) \in \Pi_B(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{H}_1$. Upon examining (2.10), we observe that lines 4-7 in Alg. 1 (i.e., the four steps of the ERA) can be replaced with the update rule

$$\mathbf{a}_{k+1} \leftarrow (P_A \circ P_B)(\mathbf{a}_k) \tag{2.11}$$

Evidently, the ERA produces a sequence in the object domain by alternating projections between A and B. All algorithms of the form described by (2.11) can be shown (Thm. A.1-

Algorithm 1 Error Reduction Algorithm

1: $k \leftarrow 0$
2: $\mathbf{a}_k \leftarrow initialize$
3: **repeat**
4: $\quad \widehat{\mathbf{a}}_k \leftarrow \mathcal{F}\mathbf{a}_k$
5: $\quad \tilde{\mathbf{b}}_k \leftarrow P_{\widetilde{B}}(\widehat{\mathbf{a}}_k)$
6: $\quad \mathbf{b}_k \leftarrow \mathcal{F}^{-1}\tilde{\mathbf{b}}_k$
7: $\quad \mathbf{a}_{k+1} \leftarrow P_A(\mathbf{b}_k)$
8: $\quad k \leftarrow k + 1$
9: **until** convergence

1) to produce a sequence $\{\mathbf{a}_k\}$ such that[3]

$$d(\mathbf{a}_{k+1}, B) \leq d(\mathbf{a}_k, B) \tag{2.12}$$

Together, results (2.9) and (2.12) imply that the ERA is guaranteed to produce a sequence with nonincreasing error – a property that is often referred to as *error reduction* [113, 124]. Since d is bounded below, the sequence $\{d(\mathbf{a}_k, B)\}_{k=0}^{\infty} \subseteq \mathbb{R}$ is guaranteed to converge. In special cases, such as when A and B are convex sets, $\{d(\mathbf{a}_k, B)\}$ can be shown to converge to a local minimum of $d(A, B)$. However, it remains an open problem to show that $\{d(\mathbf{a}_k, B)\}$ converges to a local minimum for the general case. Still, the empirical success of the ERA has made it, and its variants, a commonly employed solution for phase retrieval.

2.5 Error Reduction via Alternating Projections

The error reduction property of the ERA has only been proved for unitary $\mathcal{F}$. Thus, for the discrete waveform synthesis problem in which $\mathcal{H}_1 = \mathbb{C}^N$ and $\mathcal{H}_2 = \mathbb{C}^M$, the ERA property is proved to hold only for $N = M$. That is, error reduction is proved only when the matrix representing the Fourier transform is square. In this section, we consider the more general case of $N \neq M$. When $N \leq M$, we will denote the transform matrix by $\mathbf{U}$, and we will require only $\mathbf{U}^H\mathbf{U} = \mathbf{I}$ (left unitary). When $N \geq M$, we will denote the Fourier matrix by $\mathbf{V}$, and require only $\mathbf{V}\mathbf{V}^H = \mathbf{I}$ (right unitary). The theory will hold for any $\mathbf{U}$ or $\mathbf{V}$ that meet these requirements. In the spectrum shaping problem for radar waveform optimization, however, each $[\mathbf{U}\mathbf{x}]_m$ or $[\mathbf{V}\mathbf{x}]_m$ represents the discrete-time Fourier transform (DTFT) of some vector $\mathbf{x} \in \mathbb{C}^N$ evaluated at a particular frequency indexed by m. In App. A.2, we discuss how to choose frequency sets so that $\mathbf{U}$ and $\mathbf{V}$

[3]Levi and Stark provide a more general treatment in [124], of which this result is a direct consequence. However, for a simple proof, see Thm. A.1-1 in the Appendix.

satisfy these requirements. This amounts to building matrices from appropriately selected rows or columns from a DTFT matrix of appropriate dimensions.

2.5.1 Fewer Time Samples than Frequency Samples ($N \leq M$)

Jackson *et al.* [127] applied the ERA to the radar waveform synthesis problem in which the number of time domain samples is less than the number of optimal FTM samples. However, no analysis of convergence behavior or solution optimality was provided. Pillai *et al.* [129] considered the error reduction property of the ERA for waveform synthesis, but did not address the $N \neq M$ case. We now prove that the Gerchberg-Saxton algorithm possesses the error reduction property when $N < M$. The optimality of the solution is discussed in the following sections.

In order to apply the ERA (Alg. 1) to this problem, we must define appropriate constraint sets and projection operators. Let us define the image and object domain constraint sets by

$$A := \left\{ \mathbf{x} \in \mathbb{C}^N : |\mathbf{x}| = \mathbf{p} \right\} \tag{2.13}$$

and

$$\widetilde{B} := \left\{ \tilde{\mathbf{b}} \in \mathbb{C}^M : |\tilde{\mathbf{b}}| = \mathbf{q} \right\} \tag{2.14}$$

respectively. Further, define the mappings $P_A : \mathbb{C}^N \rightarrow A$ and $P_{\widetilde{B}} : \mathbb{C}^M \rightarrow \widetilde{B}$ by

$$\left[P_A(\mathbf{x}) \right]_n := p_n \exp\left(j \angle x_n \right) \tag{2.15}$$

$$\left[P_{\widetilde{B}}(\tilde{\mathbf{x}}) \right]_m := q_m \exp\left(j \angle \tilde{x}_m \right) \tag{2.16}$$

respectively. It can be shown that $P_A(\mathbf{x}) \in \Pi_A(\mathbf{x})$ and $P_{\widetilde{B}}(\tilde{\mathbf{x}}) \in \Pi_{\widetilde{B}}(\tilde{\mathbf{x}})$ for all $\mathbf{x} \in \mathbb{C}^N$ and $\tilde{\mathbf{x}} \in \mathbb{C}^M$ (Thm. A.1-2, A.1-3). The relationship between these sets is illustrated in Fig. 2.6.

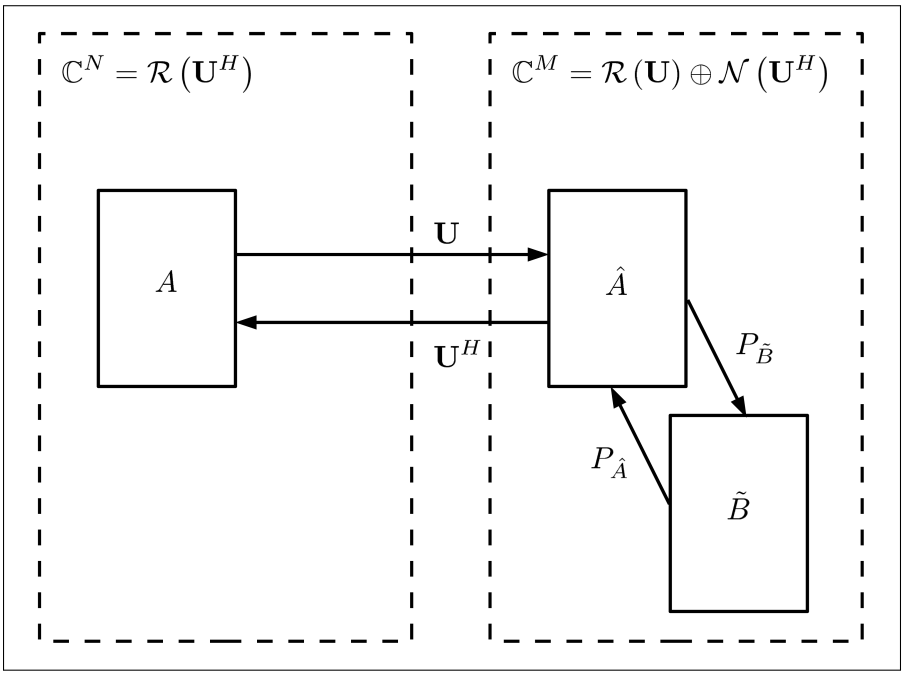


Figure 2.6: Relevant sets in the object and image domains when $N \leq M$.

This particular form of the ERA is known as the Gerchberg-Saxton Algorithm [113]. The GSA is given in Alg. 2 where lines 3-6 correspond to lines 4-7 in Alg. 1.

The error reduction property of the ERA was proved only for unitary $\mathbf{U}$ (i.e., when $N = M$). In order to prove the error reduction property when $N < M$, we now consider its alternating projections nature. Denote the image of set A under $\mathbf{U}$ by

$$\widehat{A} := \mathbf{U}(A) \tag{2.17}$$

Algorithm 2 GSA for $N \leq M$

1: $\mathbf{a} \leftarrow initialize$
2: **repeat**
3: $\quad \widehat{\mathbf{a}} \leftarrow \mathbf{U}\mathbf{a}$
4: $\quad \tilde{\mathbf{b}} \leftarrow P_{\tilde{B}}(\widehat{\mathbf{a}})$
5: $\quad \mathbf{b} \leftarrow \mathbf{U}^H\tilde{\mathbf{b}}$
6: $\quad \mathbf{a} \leftarrow P_A(\mathbf{b})$
7: **until** convergence

and define the mapping $P_{\widehat{A}} : \mathbb{C}^M \rightarrow \widehat{A}$ to be

$$P_{\widehat{A}}(\tilde{\mathbf{x}}) := \mathbf{U} P_A \left(\mathbf{U}^H \tilde{\mathbf{x}} \right) \tag{2.18}$$

It can be shown (Thm. A.1-4) that $P_{\widehat{A}}(\tilde{\mathbf{x}}) \in \Pi_{\widehat{A}}(\tilde{\mathbf{x}})$ for all $\tilde{\mathbf{x}} \in \mathbb{C}^M$. Using these definitions, Alg. 2 can be written in the equivalent alternating projections form given by Alg. 3. As an AP algorithm, Alg. 3 has the property that $d(\widehat{\mathbf{a}}_{k+1}, \widetilde{B}) \leq d(\widehat{\mathbf{a}}_k, \widetilde{B})$. Furthermore, it can be shown (Thm. A.1-6) that this point-set distance is equal to the distance between the time domain vector's FTM and the desired FTM. Thus,

$$d^2(\widehat{\mathbf{a}}, \widetilde{B}) = \left| \left| |\widehat{\mathbf{a}}| - \mathbf{q} \right| \right|^2 \tag{2.19}$$

From this relationship, we can see why the GSA is an effective algorithm. The nonincreasing point set distance is a measure of the mean square error between the optimal FTM and the FTM of the current estimate $|\widehat{\mathbf{a}}_k|$. As a practical matter, the algorithm can be initialized by setting each entry of the initial phase function $\angle \mathbf{x}_n$ equal to an i.i.d. random value distributed uniformly over $[0, 2\pi)$, and program convergence can be determined by monitoring the change in $d(\widehat{\mathbf{a}}, \widetilde{B})$.

2.5.2 More Time Samples than Frequency Samples ($N \geq M$)

Jackson *et al.* [127] state that the ERA cannot be applied when the number of samples in the time domain is greater than or equal to the number of samples in the frequency domain,

Algorithm 3 Alternating Projections for $N \leq M$

1: $\mathbf{a} \leftarrow initialize$
2: $\widehat{\mathbf{a}} \leftarrow \mathbf{U}\mathbf{a}$
3: **repeat**
4: $\quad \widehat{\mathbf{a}} \leftarrow P_{\widehat{A}}(P_{\widetilde{B}}(\widehat{\mathbf{a}}))$
5: **until** convergence
6: $\mathbf{a} \leftarrow \mathbf{U}^H \widehat{\mathbf{a}}$

but an alternative algorithm is not provided. We take an alternating projections approach to develop a new error reducing algorithm, which cannot be shown to be a special case of the ERA. We begin by noting that if we simply replace $\mathbf{U}$ with $\mathbf{V}$, as is done in Alg. 4, then the resulting algorithm does not possess the error reduction property. This is because

$$P_{\widehat{A}}(\tilde{\mathbf{x}}) := \mathbf{V} P_A\left(\mathbf{V}^H \tilde{\mathbf{x}}\right) \tag{2.20}$$

is not a member of the projection $\Pi_{\widehat{A}}(\tilde{\mathbf{x}})$.

For $N \geq M$, consider the set of all time domain vectors whose frequency domain representation have the desired modulus, which we will denote

$$B := \left\{\mathbf{b} \in \mathbb{C}^N : |\mathbf{V}\mathbf{b}| = \mathbf{q}\right\} \tag{2.21}$$

The relevant sets for the $N \geq M$ case are shown in Fig. 2.7. Next, define the mapping $P_B : \mathbb{C}^N \rightarrow B$ by

$$P_B(\mathbf{x}) := \mathbf{V}^H P_{\widetilde{B}}(\mathbf{V}\mathbf{x}) + (\mathbf{I} - \mathbf{V}^H\mathbf{V})\mathbf{x} \tag{2.22}$$

It can be shown (Thm. A.1-5) that $P_B(\mathbf{x}) \in \Pi_B(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{C}^N$. We can then write the AP algorithm as listed in Alg. 5. Due to its AP nature, this algorithm possesses the property that $d(\mathbf{a}_{k+1}, B) \leq d(\mathbf{a}_k, B)$. Furthermore, it can be shown (Thm. A.1-7) that $d(\mathbf{a}, B) = ||\,|\widehat{\mathbf{a}}| - \mathbf{q}\,||$ for all $\mathbf{a} \in A$. Thus, this algorithm has the same meaningful error

Algorithm 4 GSA for $N \geq M$

1: $\mathbf{a} \leftarrow initialize$
2: **repeat**
3: $\quad \widehat{\mathbf{a}} \leftarrow \mathbf{V}\mathbf{a}$
4: $\quad \tilde{\mathbf{b}} \leftarrow P_{\widetilde{B}}(\widehat{\mathbf{a}})$
5: $\quad \mathbf{b} \leftarrow \mathbf{V}^H\tilde{\mathbf{b}}$
6: $\quad \mathbf{a} \leftarrow P_A(\mathbf{b})$
7: **until** convergence

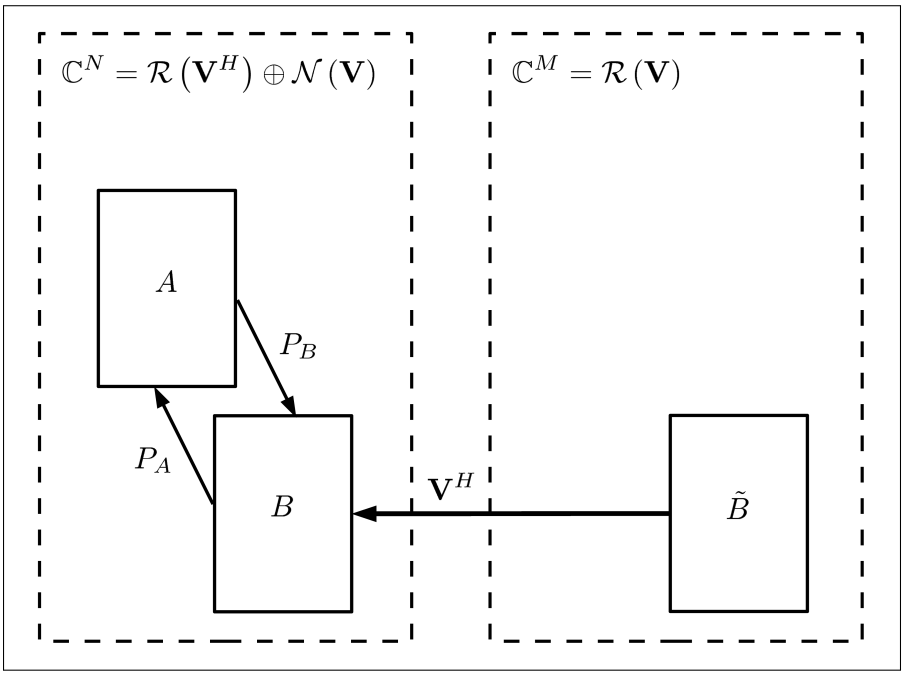


Figure 2.7: Relevant sets in the object and image domains when $N \geq M$.

reduction property as in the case of $N \leq M$.

Alg. 5 can be written in the equivalent form shown in Alg. 6. This algorithm is similar to the naive GSA formulation in Alg. 4. However, the two algorithms are not equivalent due to the null space component $(\mathbf{I} - \mathbf{V}^H\mathbf{V})\mathbf{a}$ in line 5. We note that when $N \leq M$, we can recover the GSA (Alg. 2) from Alg. 6 by replacing $\mathbf{V}$ with $\mathbf{U}$ and observing that $\mathbf{U}^H\widehat{\mathbf{a}} = \mathbf{U}^H\mathbf{U}\mathbf{a} = \mathbf{a}$. Thus, the GSA can be seen as a special case of this alternating projections algorithm when $N \leq M$.

Algorithm 5 Alternating Projections for $N \geq M$

1: $\mathbf{a} \leftarrow initialize$
2: **repeat**
3: $\quad \mathbf{a} \leftarrow P_A(P_B(\mathbf{a}))$
4: **until** convergence

Algorithm 6 Modified GSA for $N \geq M$

$\mathbf{a} \leftarrow initialize$
repeat
  $\widehat{\mathbf{a}} \leftarrow \mathbf{V}\mathbf{a}$
  $\tilde{\mathbf{b}} \leftarrow P_{\widetilde{B}}(\widehat{\mathbf{a}})$
  $\mathbf{b} \leftarrow \mathbf{a} - \mathbf{V}^H\widehat{\mathbf{a}} + \mathbf{V}^H\tilde{\mathbf{b}}$
  $\mathbf{a} \leftarrow P_A(\mathbf{b})$
until convergence

2.6 Nonlinear Programming

2.6.1 Problem Formulation

The discrete-time spectrum shaping problem with which we are interested can be stated as follows: *Given* $\mathbf{p} \in \mathbb{R}_+^N$ *and* $\mathbf{q} \in \mathbb{R}_+^M$*, find an* $\mathbf{x} \in \mathbb{C}^N$ *with* $|\mathbf{x}| = \mathbf{p}$ *such that* $|\widehat{\mathbf{x}}|$ *is of minimal Euclidean distance to* $\mathbf{q}$. For the waveform optimization problem, $\mathbf{p}$ represents the desired time domain envelope, and $\mathbf{q}$ represents the desired FTM. If we define the cost function $J : \mathbb{C}^N \rightarrow \mathbb{R}$ to be

$$J(\mathbf{x}) := \left|\left||\mathbf{x}| - \mathbf{q}\right|\right|^2 \tag{2.23}$$

then the discrete-time spectrum shaping problem becomes that of solving the optimization problem

$$\inf_{\mathbf{a} \in A} J(\mathbf{a}) \tag{2.24}$$

where A is defined in (2.13). Each entry of the vector $\mathbf{a}$ is defined by $a_n = p_n e^{j\phi_n}$ where $\phi_n \in \mathbb{R}$ for all n. Thus, the constrained optimization problem in (2.24) can be converted into an unconstrained problem by minimizing J with respect to $\boldsymbol{\phi} = [\phi_1 \ldots \phi_N]^T$ instead of $\mathbf{a}$. A solution can be obtained by applying standard direct search methods such as gradient descent or the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method [130]. The gradient of the cost function J with respect to the phase vector is provided in App. A.3.

2.6.2 GD vs. AP

Fienup [113] showed the similarity between gradient descent and the GSA when $N = M$. We continue that analysis for $N \leq M$. The gradient descent and alternating projections algorithms for the waveform synthesis problem are listed in full detail in Alg. 7 and Alg. 8, respectively. Notice that for $N \leq M$, the algorithms differ only in the phase update (line 12). For this case, when k is large such that the gradient descent algorithm is converging, and thus, the phase is changing slowly, we can express the n^{th} element of the updated phase by

$$\phi_n - \alpha \mathrm{Im}\left\{a_n b_n^*\right\} = \phi_n - \alpha \left|p_n\right| \left|b_n^*\right| \sin\left(\phi_n - \angle b_n\right) \tag{2.25}$$

$$\approxeq \phi_n - \alpha \left|p_n\right| \left|b_n^*\right| \left[\phi_n - \angle b_n\right] \tag{2.26}$$

$$= \angle b_n + \left(1 - \alpha \left|p_n\right| \left|b_n^*\right|\right) \phi_n \tag{2.27}$$

Thus, when the phase is slowly changing, we can observe that the GSA behaves like gradient descent with variable step size

$$\alpha_n = \frac{1}{\left|p_n\right| \left|b_n^*\right|} \tag{2.28}$$

This would suggest that the GSA converges to a local minimizer of J for the case of $N \leq M$, though this conjecture has not been proved. This correspondence between the GSA and gradient descent may help to explain the occurrence of what Levi and Stark call *traps* and *tunnels* in the GSA [124]. Traps may correspond to local minima, whereas tunnels may occur when J has long narrow valleys. The reader is referred to [124] for more detail.

The convergence behavior of the GD algorithm is very sensitive to the selection of the step size α. In practice, there is no means of determining α *a priori*, and experience has shown that the GSA can converge much faster than GD for poorly selected α. To remove this sensitivity to step size, and to improve convergence time, one may implement a quasi-

Algorithm 7 GD with Fixed Step Size

$\boldsymbol{\phi} \leftarrow$ uniform random over $[0, 2\pi)$
$\mathbf{a} \leftarrow \mathbf{p} \odot \exp(j\boldsymbol{\phi})$
$\widehat{\mathbf{a}} \leftarrow \mathbf{F}\mathbf{a}$
$J_{new} \leftarrow \left|\left|\left|\widehat{\mathbf{a}}\right| - \mathbf{q}\right|\right|^2$
repeat
  $\tilde{\mathbf{b}} \leftarrow \mathbf{q} \odot \exp(j\angle\widehat{\mathbf{a}})$
  if $N \leq M$ **then**
    $\mathbf{b} \leftarrow \mathbf{F}^H\tilde{\mathbf{b}}$
  else
    $\mathbf{b} \leftarrow \mathbf{F}^H(\tilde{\mathbf{b}} - \widehat{\mathbf{a}})$
  end if
  $\boldsymbol{\phi} \leftarrow \boldsymbol{\phi} - \alpha\mathrm{Im}\{\mathbf{a} \odot \mathbf{b}^*\}$
  $\mathbf{a} \leftarrow \mathbf{p} \odot \exp(j\boldsymbol{\phi})$
  $\widehat{\mathbf{a}} \leftarrow \mathbf{F}\mathbf{a}$
  $J_{old} \leftarrow J_{new}$
  $J_{new} \leftarrow \left|\left|\left|\widehat{\mathbf{a}}\right| - \mathbf{q}\right|\right|^2$
until $|J_{old} - J_{new}| < \epsilon$

Algorithm 8 Alternating Projections

$\boldsymbol{\phi} \leftarrow$ uniform random over $[0, 2\pi)$
$\mathbf{a} \leftarrow \mathbf{p} \odot \exp(j\boldsymbol{\phi})$
$\widehat{\mathbf{a}} \leftarrow \mathbf{F}\mathbf{a}$
$J_{new} \leftarrow \left|\left|\left|\widehat{\mathbf{a}}\right| - \mathbf{q}\right|\right|^2$
repeat
  $\tilde{\mathbf{b}} \leftarrow \mathbf{q} \odot \exp(j\angle\widehat{\mathbf{a}})$
  if $N \leq M$ **then**
    $\mathbf{b} \leftarrow \mathbf{F}^H\tilde{\mathbf{b}}$
  else
    $\mathbf{b} \leftarrow \mathbf{a} + \mathbf{F}^H(\tilde{\mathbf{b}} - \widehat{\mathbf{a}})$
  end if
  $\boldsymbol{\phi} \leftarrow \angle\mathbf{b}$
  $\mathbf{a} \leftarrow \mathbf{p} \odot \exp(j\boldsymbol{\phi})$
  $\widehat{\mathbf{a}} \leftarrow \mathbf{F}\mathbf{a}$
  $J_{old} \leftarrow J_{new}$
  $J_{new} \leftarrow \left|\left|\left|\widehat{\mathbf{a}}\right| - \mathbf{q}\right|\right|^2$
until $|J_{old} - J_{new}| < \epsilon$

Newton method with inexact line search [131]. It should be noted that inexact line search methods do require user parameters. However, algorithm behavior is generally insensitive to the choice of these parameters.

2.6.3 The PSD Synthesis Problem

For some waveform optimization problems, it may be more desirable to find a waveform of the appropriate duration that possesses the desired envelope in the time domain and a power spectral density (PSD) that is as close as possible to the square of the optimal FTM (e.g., [28, 112]). We shall refer to this problem as the *PSD synthesis* problem. Let

$$\widetilde{B}_2 := \left\{\tilde{\mathbf{x}} \in \mathbb{C}^M : |\tilde{\mathbf{x}}|^2 = \mathbf{q}^2\right\} \tag{2.29}$$

denote the set of waveforms satisfying the PSD requirement. Clearly, $\widetilde{B}_2 = \widetilde{B}$. Thus, the alternating projections approach can be used to find an $\mathbf{a} \in A$ such that $\widehat{\mathbf{a}}$ minimizes

the distance from $\widehat{A}$ to $\widetilde{B}_2$. However, this point-set distance is not equivalent to the cost function that is a true measure of error, namely, $J_2 : \mathbb{C}^N \to \mathbb{R}$ defined by

$$J_2(\mathbf{x}) := \left|\left||\mathbf{x}|^2 - \mathbf{q}^2\right|\right|^2 \tag{2.30}$$

As such, the alternating projections algorithm is not guaranteed to produce iterates such that $J_2(\mathbf{a}_{k+1}) \leq J_2(\mathbf{a}_k)$. However, J and J_2 are closely related, and it is not unreasonable to expect the algorithm to produce useful waveforms. As before, we can formulate an unconstrained nonlinear programming problem by minimizing J_2 with respect to the signal phase, and we can apply standard descent algorithms. (The gradient of J_2 with respect to the signal phase is provided in App. A.3.) We provide numerical examples comparing the AP approach to descent methods in the following section.

2.7 Examples

2.7.1 Overview

In this section, we demonstrate the efficacy of the AP approach by solving example spectrum shaping and PSD synthesis problems using the AP algorithm, gradient descent with fixed step size, and the BFGS quasi-Newton method with inexact line search described in [131]. In each example, the time domain waveform is specified to have constant modulus (i.e., $\mathbf{p} = \mathbf{1}$), and the optimal FTM (i.e., $\mathbf{q}$) is taken from [28]. The algorithms were run 100,000 times for each example, with each trial corresponding to a new random phase initialization as described in Sec. 2.5. All algorithms were given the same initialization at the beginning of each trial. The average convergence behavior of each algorithm was assessed by averaging the normalized cost at each iteration over all trials. By *normalized cost*, we

mean the reduction in cost from the initial guess as measured by

$$\overline{J}(\mathbf{a}_k) := \frac{J(\mathbf{a}_k)}{J(\mathbf{a}_0)} \text{ or } \overline{J}_2(\mathbf{a}_k) := \frac{J_2(\mathbf{a}_k)}{J_2(\mathbf{a}_0)} \tag{2.31}$$

depending upon the problem. The probability density function (pdf) of the normalized cost at termination was also estimated by computing the normalized histogram. Note that the gradient descent step size was determined manually for each example in order to produce good convergence behavior.

2.7.2 Spectrum Shaping: $N < M$

Consider the problem of synthesizing a length $N = 32$ constant modulus waveform possessing an FTM close to the length $M = 128$ vector $\mathbf{q}$ shown in Fig. 2.8. The average normalized cost as a function of iteration number is shown in Fig. 2.9 for the BFGS quasi-Newton method with inexact line search (QN), the gradient descent algorithm (GD), and the GSA. As expected, the quasi-Newton method converged to a solution in fewer iterations than either the GD or GSA. However, all methods produced solutions of nearly identical quality. This is apparent in Fig. 2.10, which shows the estimated pdf of the final normalized cost. We emphasize that, while the performance and efficiency of the GD and GSA were comparable, the GD step size was manually tuned to achieve this result. The quasi-Newton method requires fewer iterations to converge to a solution, but each QN iteration requires many more calculations than an iteration of GD or GSA. To illustrate this point, the simulation was repeated with the exception that each algorithm was allowed to terminate a given trial if the change in the cost function from time k to $k + 1$ was less than 0.15% for five consecutive iterations. This stopping criteria was chosen so as to maintain the quality of the final solution. With this stopping criteria, the QN method required an average of 40% more CPU time than the GSA.

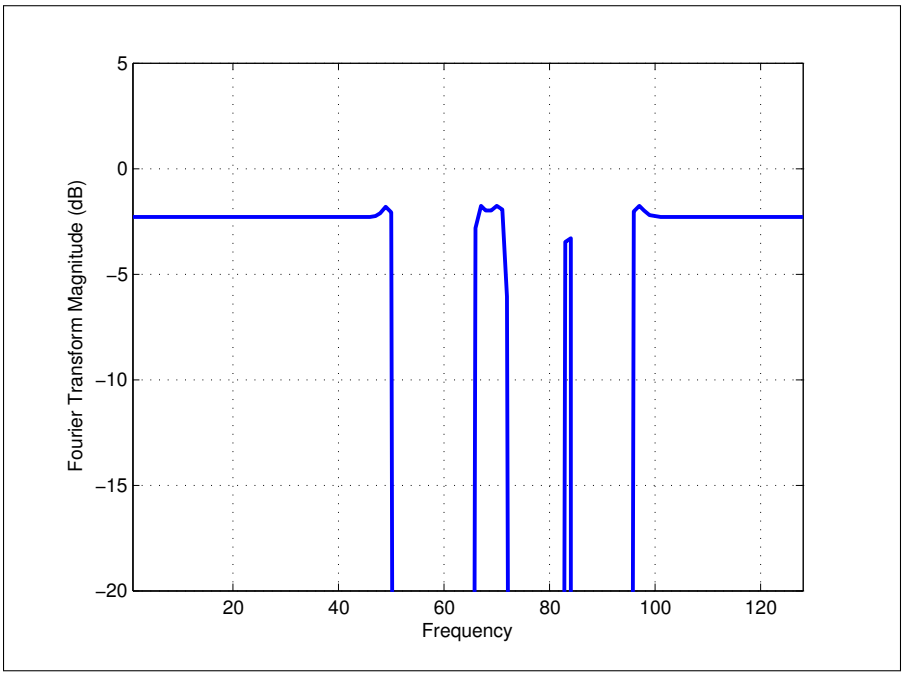


Figure 2.8: Desired FTM ($\mathbf{q}$) for the $N < M$ example corresponding to $N = 32$, $L = 4$, $M = 128$ and $\Gamma = \{0, 1, \ldots, 127\}$.

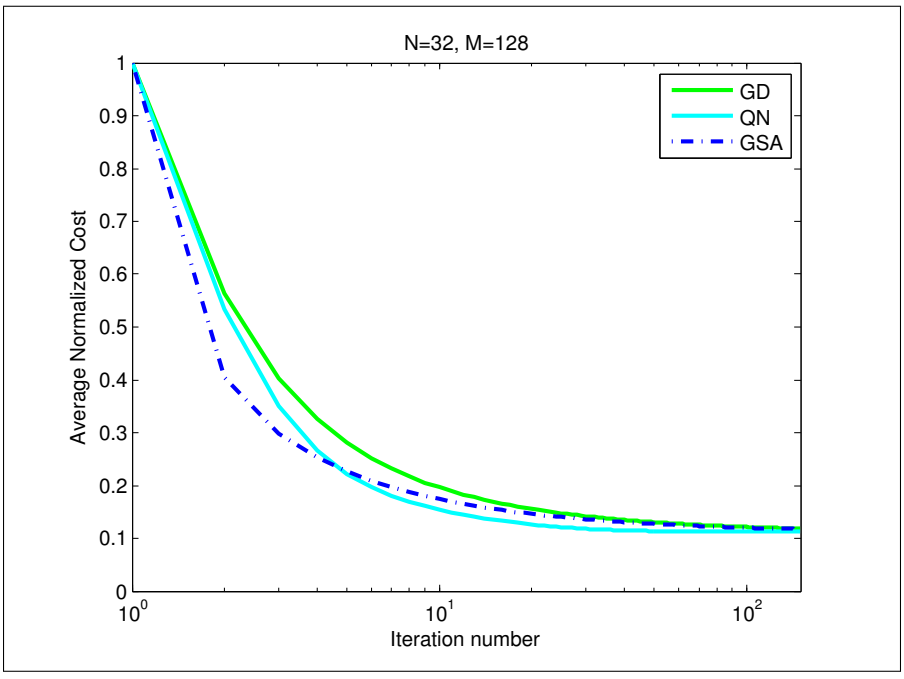


Figure 2.9: Average normalized cost vs. iteration number for the $N < M$ spectrum shaping example. GD step size equal to 0.5.

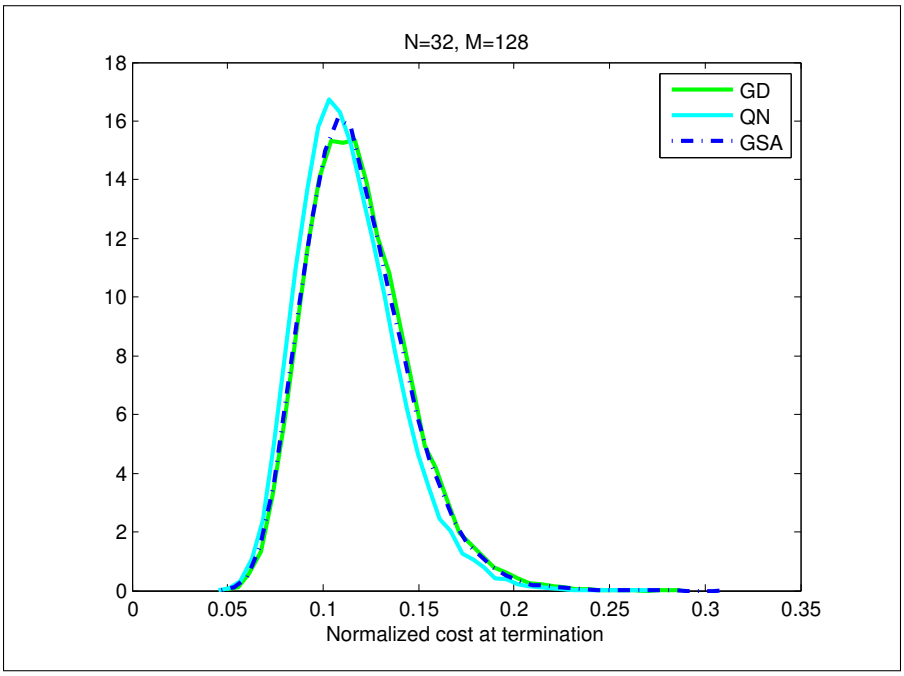


Figure 2.10: Estimated pdf of normalized cost at termination for the $N < M$ spectrum shaping example. GD step size equal to 0.5.

2.7.3 Spectrum Shaping: $N > M$

Consider the problem of synthesizing a length $N = 128$ constant modulus waveform possessing an FTM close to the length $M = 32$ vector $\mathbf{q}$ shown in Fig. 2.11. The applicable algorithms for this problem are QN, GD, and AP (Alg. 5). However, we also consider the naive GSA (Alg. 4) to highlight its difference from AP when $N > M$. The average normalized cost as a function of iteration number is shown in Fig. 2.12, and the final cost histograms are shown in Fig. 2.13. As expected, AP performs better than the naive GSA, with average performance nearly equivalent to GD with a manually-tuned step size. Again, the QN method performs better than the rest in terms of number of iterations to convergence and quality of final solution. However, the quality of final solution may be somewhat exaggerated in Fig. 2.13, as the QN method provided only an additional 0.05% reduction from the initial cost. However, the difference in run time was more significant than before. When the simulation was repeated using the termination criteria previously described, the QN required an average of 800% more CPU time than the AP algorithm.

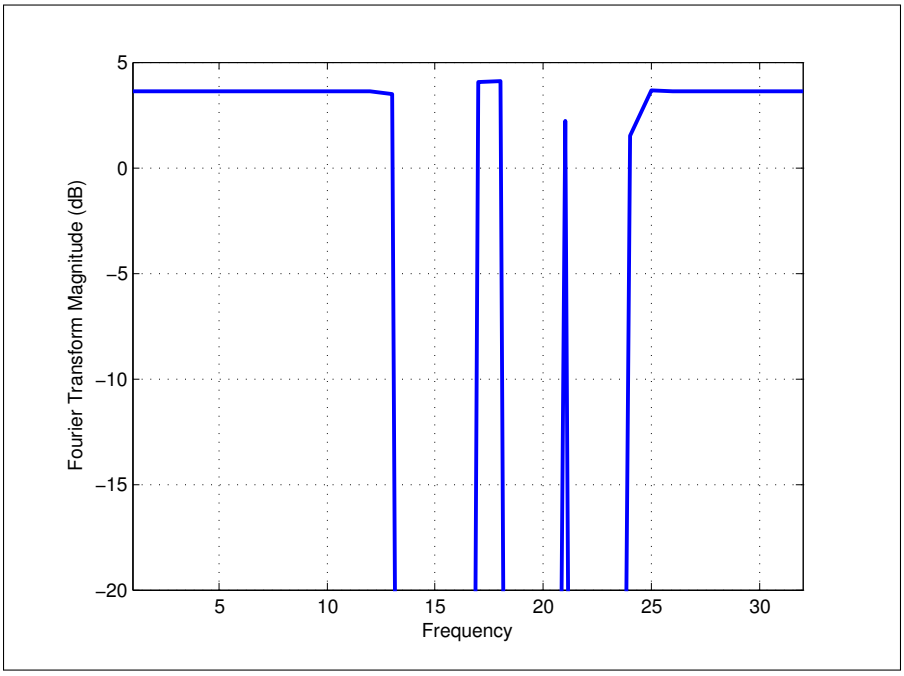


Figure 2.11: Desired FTM ($\mathbf{q}$) for the $N > M$ example corresponding to $N = 128$, $L = 1/4$, $M = 32$ and $\Gamma = \{0, 1, \ldots, 31\}$.

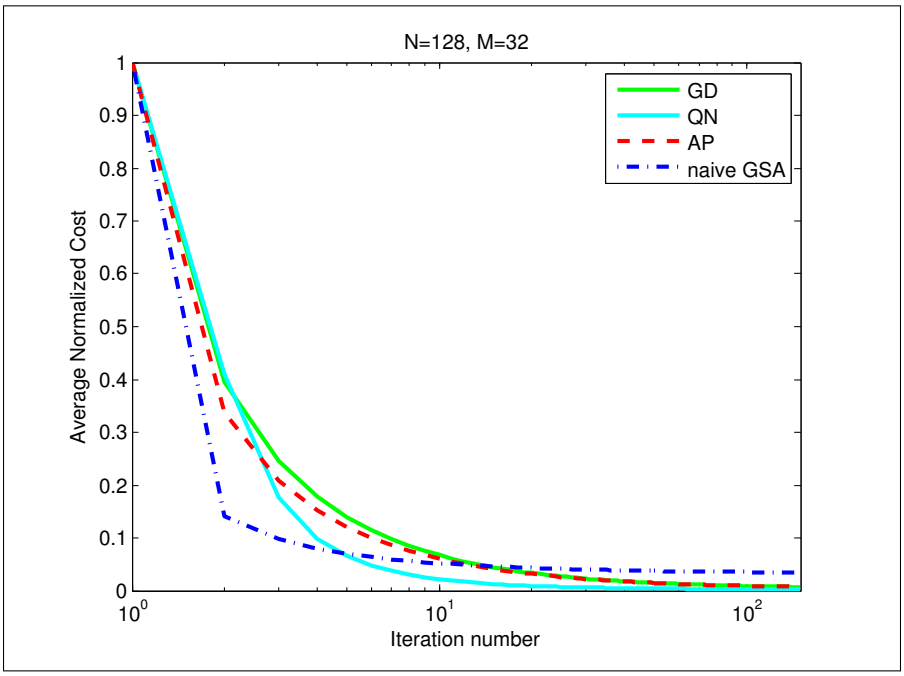


Figure 2.12: Average normalized cost vs. iteration number for the $N > M$ spectrum shaping example. GD step size equal to 0.5.

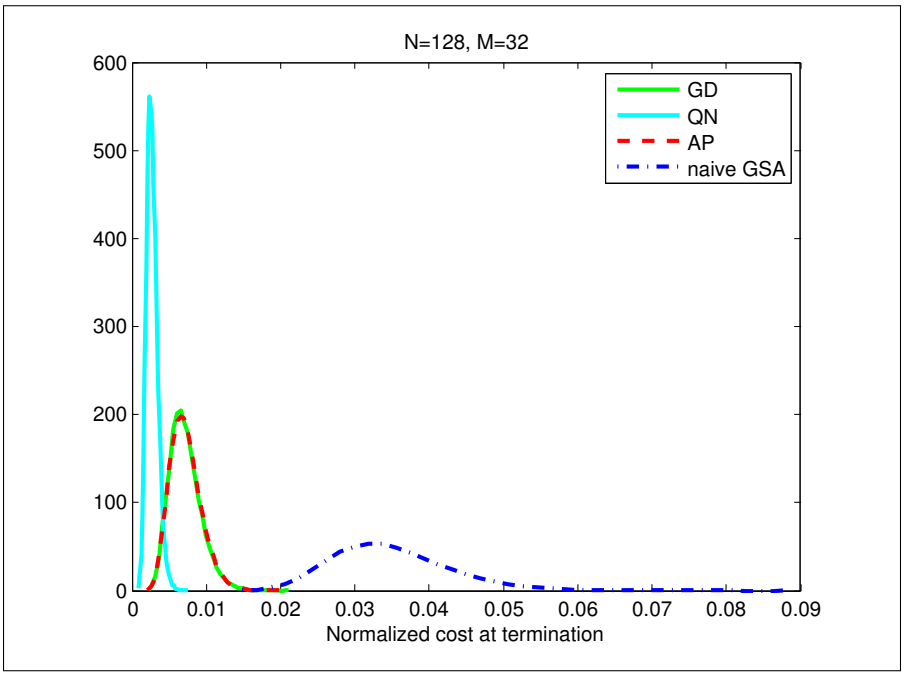


Figure 2.13: Estimated pdf of normalized cost at termination for the $N > M$ spectrum shaping example. GD step size equal to 0.5.

2.7.4 PSD Synthesis: $N < M$

Consider the problem of synthesizing a length $N = 32$ constant modulus waveform possessing a power spectral density close to the length $M = 128$ vector $\mathbf{q}$, the square root of which is shown (in dB) in Fig. 2.8. As discussed in Sec. 2.6.3, the GSA is not guaranteed to reduce the true error J_2, and gradient methods are expected to perform better. This is demonstrated in Figs. 2.14 and Fig. 2.15. It is notable that the GSA does substantially reduce the normalized cost.

2.7.5 PSD Synthesis: $N > M$

Consider the problem of synthesizing a length $N = 128$ constant modulus waveform possessing a power spectral density close to the length $M = 32$ vector $\mathbf{q}$, the square root of which is shown (in dB) in Fig. 2.11. Again, we consider GD, AP (Alg. 5), and the naive GSA (Alg. 4) for comparison purposes. The average normalized cost as a function of iteration number is shown in Fig. 2.16, and the final cost histograms are shown in Fig. 2.17. As

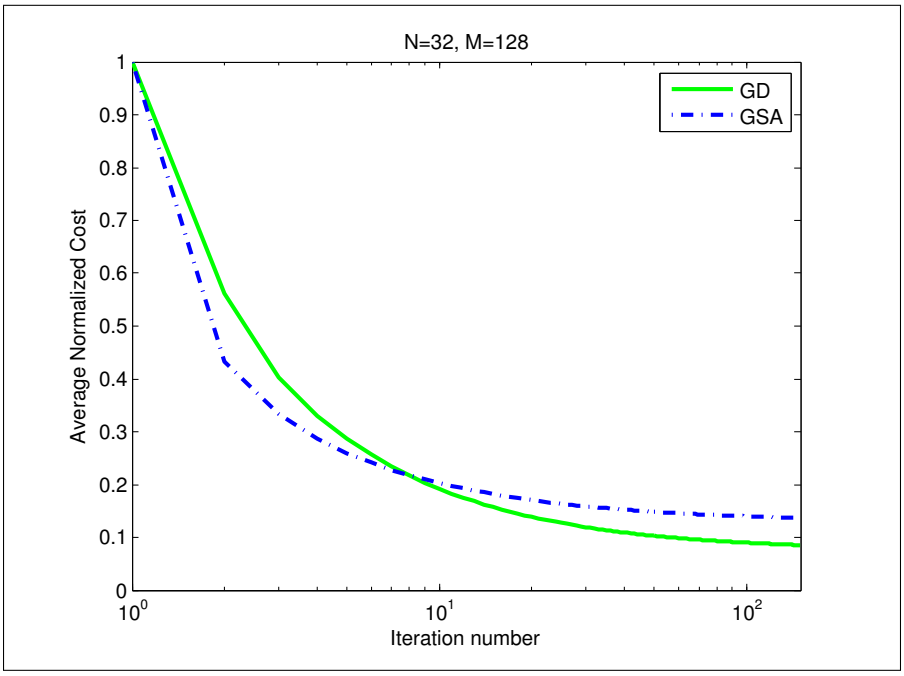


Figure 2.14: Average normalized cost vs. iteration number for the $N < M$ PSD synthesis example. GD step size equal to 0.5.

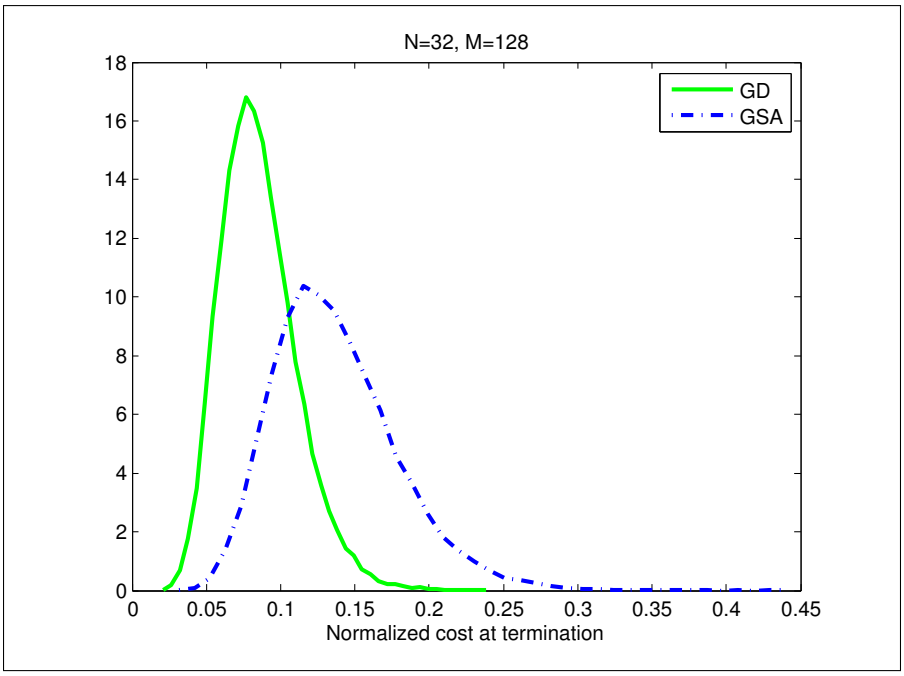


Figure 2.15: Estimated pdf of normalized cost at termination for the $N < M$ PSD synthesis example. GD step size equal to 0.5.

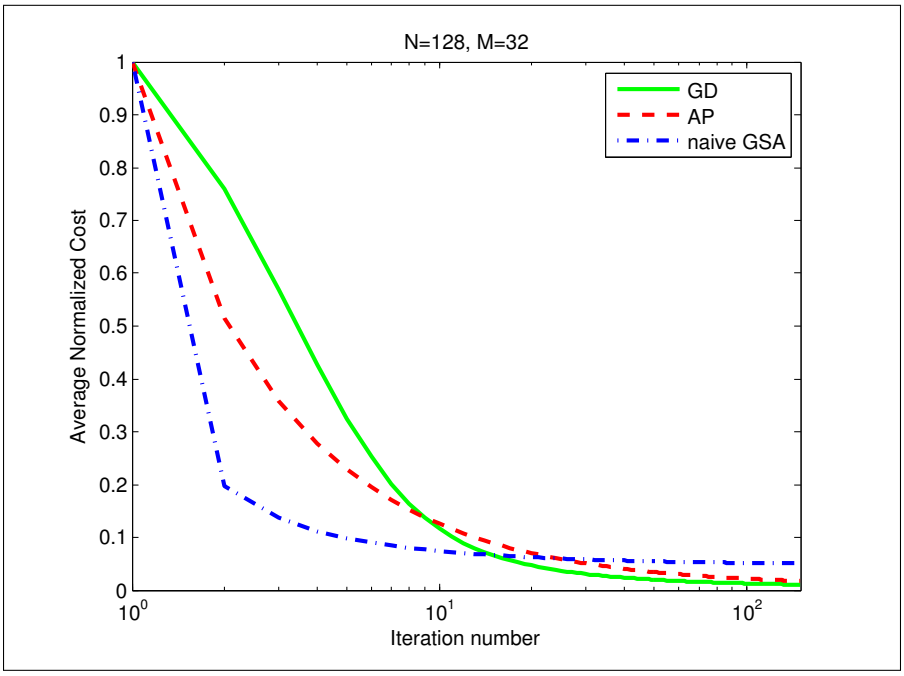


Figure 2.16: Average normalized cost vs. iteration number for the $N > M$ PSD synthesis example. GD step size equal to 0.05.

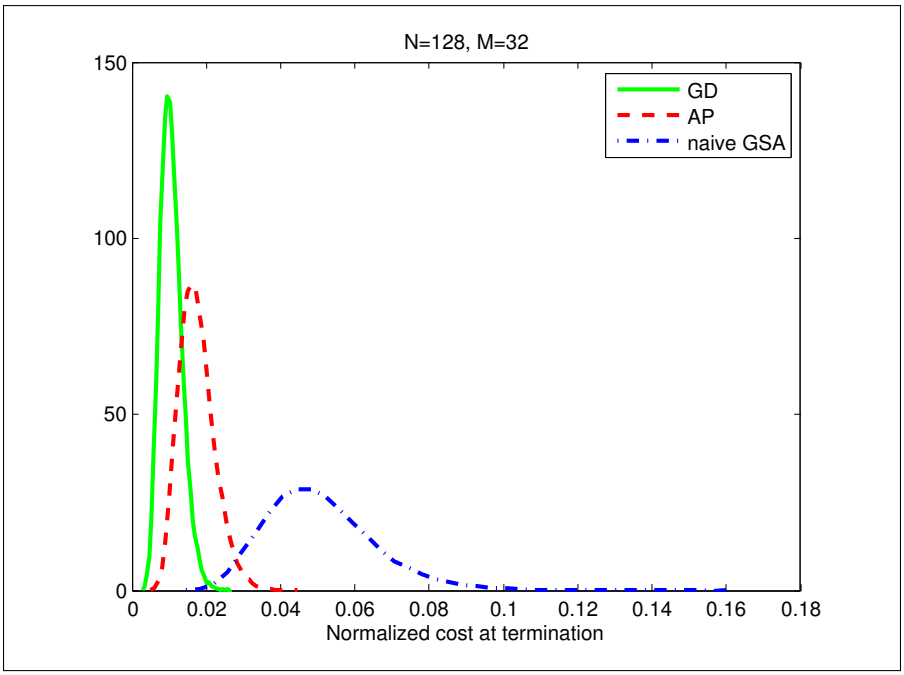


Figure 2.17: Estimated pdf of normalized cost at termination for the $N > M$ PSD synthesis example. GD step size equal to 0.05.

expected, the GD algorithm has superior average convergence and performance properties. However, AP substantially reduces the normalized cost, and performs nearly as well as the GD algorithm. Also, as expected, the naive GSA reduces the normalized cost, but performs worse than either of the other two algorithms.

2.8 Conclusion

We have considered the waveform synthesis problem found in radar waveform optimization, and showed that for the spectrum shaping problem the alternating projections approach can be a convenient solution in that it is (a) parameter free, (b) easier to implement than "parameter free" gradient methods, and (c) it can provide near optimal solutions with lower computational cost. While we have proved its error reduction property, showing convergence to a local minimum remains an open research problem.

The PSD synthesis problem appears frequently in the radar waveform optimization literature. This problem is related to the spectrum shaping problem, but the two problems are not equivalent. As such the AP algorithm is expected to provide a suboptimal solution. This was confirmed by two numerical examples. However, the difference between the optimal GD algorithm and the suboptimal AP algorithm was not extreme in either example. As such, the relative accuracy of solution, lack of user specified parameters, and computational speed make the AP an applicable algorithm for the PSD synthesis problem as well.

Chapter 3

Waveform-Optimized Performance and Ambiguity Function Constraints

In the previous chapter, we showed that modulus constraints could be addressed by solving a secondary optimization problem. However, for many applications the cross-ambiguity function of the transmit signal and the receive filter must also be constrained. This occurs, for example, when a radar must detect an unknown number of targets in colored noise. In Sec. 3.1, we review the well studied problem of detecting a known signal in noise so that the idea of waveform-optimized performance can be introduced. In Sec. 3.2, we formulate the waveform optimization problem for various receiver architectures in the multiple unknown targets scenario, and we use the idea of waveform-optimized performance to anticipate the efficacy of each approach. In Sec. 3.3, we present numerical examples that demonstrate differences between the waveform-optimized performance of each receive architecture. Concluding remarks for this chapter are given in Sec. 3.4.

3.1 Waveform-Optimized Performance

In this section, we use the well studied problem of detecting a known signal in additive Gaussian noise to illustrate the idea of waveform-optimized performance. We emphasize

that the results of Sec. 3.1.1 and 3.1.3 are not new. However, reviewing these results allows us to easily formulate the joint signal/filter design problem in Sec. 3.1.2, and introduce waveform-optimized performance in Sec. 3.1.4. The results of this section also allow for a straightforward development of the multiple unknown targets scenario in Sec. 3.2.

3.1.1 Detecting a Known Signal

Consider the problem of detecting a deterministic real-valued signal in the presence of an additive wide-sense stationary (WSS) Gaussian random process. Assume the signal is zero for $t \notin [0, T]$, and that it contains negligible energy at frequencies $f \notin (f_c - B/2, f_c + B/2)$, where f_c is the center frequency, and B is the bandwidth. This detection problem reduces to choosing between the two hypotheses

$$\mathcal{H}_0 : x(t) = w(t) \tag{3.1}$$

$$\mathcal{H}_1 : x(t) = s(t) + w(t) \tag{3.2}$$

where $x(t)$ is the observed signal, $s(t)$ is the signal of interest, and $w(t)$ is the Gaussian interference.

It is often convenient to consider the case in which the complex envelope of the observed signal has been recovered and time sampled. Therefore, assume the received signal has been modulated to baseband from f_c, and that both the in-phase and quadrature channels have been sampled over the interval $[0, T]$ at a rate of $f_s = B$ Hz. The hypotheses then become

$$\mathcal{H}_0 : \mathbf{x} = \mathbf{w} \tag{3.3}$$

$$\mathcal{H}_1 : \mathbf{x} = \mathbf{s} + \mathbf{w} \tag{3.4}$$

where $\mathbf{x}, \mathbf{s}, \mathbf{w} \in \mathbb{C}^N$. Since $w(t)$ is a WSS bandpass Gaussian random process, the discrete-

time complex envelope vector $\mathbf{w}$ has a complex multivariate Gaussian probability density function (pdf) [132, pp. 513-515]. This is denoted by

$$\mathbf{w} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{K}) \tag{3.5}$$

where the mean $\boldsymbol{\mu} \in \mathbb{C}^N$ and covariance matrix $\mathbf{K} \in \mathbb{C}^{N \times N}$ are assumed known, and $\mathbf{K}$ is assumed to be positive definite. For now, we shall assume only that the decision between (3.3) and (3.4) is made by comparing the real part of a linear combination of the observations to a known threshold. That is, the test is decided according to

$$\mathrm{Re}\left\{\mathbf{h}^H \mathbf{x}\right\} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{3.6}$$

where $\mathbf{h} \in \mathbb{C}^N$ is the receive filter, and γ is chosen to maintain a specified probability of false alarm (P_{FA}). While this test statistic may seem arbitrary at first, we note that this is the form of the two most widely studied detectors: the Neyman-Pearson detector and the matched filter. Both of these detectors will be discussed subsequently, but for now we focus on analyzing the general form given by (3.6).

Under either hypothesis, $\mathbf{x}$ is an affine transformation of a complex Gaussian random vector, thus [132, pp. 505-507]

$$\mathbf{x} \sim \begin{cases} \mathcal{CN}(\boldsymbol{\mu}, \mathbf{K}) & \text{under } \mathcal{H}_0 \\ \mathcal{CN}(\mathbf{s} + \boldsymbol{\mu}, \mathbf{K}) & \text{under } \mathcal{H}_1 \end{cases}. \tag{3.7}$$

We also note that $\mathbf{h}^H \mathbf{x}$ represents a linear transformation applied to a complex Gaussian random vector, and therefore [132, pp. 505-507]

$$\mathbf{h}^H \mathbf{x} \sim \begin{cases} \mathcal{CN}\left(\mathbf{h}^H \boldsymbol{\mu}, \mathbf{h}^H \mathbf{K} \mathbf{h}\right) & \text{under } \mathcal{H}_0 \\ \mathcal{CN}\left(\mathbf{h}^H (\mathbf{s} + \boldsymbol{\mu}), \mathbf{h}^H \mathbf{K} \mathbf{h}\right) & \text{under } \mathcal{H}_1 \end{cases}. \tag{3.8}$$

Finally, the fact that $\mathbf{h}^H\mathbf{x}$ is a complex random variable implies

$$\mathrm{Re}\left\{\mathbf{h}^H\mathbf{x}\right\} \sim \begin{cases} \mathcal{N}\left(\mathrm{Re}\left\{\mathbf{h}^H\boldsymbol{\mu}\right\}, \frac{\mathbf{h}^H\mathbf{K}\mathbf{h}}{2}\right) & \text{under } \mathcal{H}_0 \\ \mathcal{N}\left(\mathrm{Re}\left\{\mathbf{h}^H(\mathbf{s}+\boldsymbol{\mu})\right\}, \frac{\mathbf{h}^H\mathbf{K}\mathbf{h}}{2}\right) & \text{under } \mathcal{H}_1 \end{cases}. \tag{3.9}$$

An inspection of (3.6) and (3.9) reveals that if we consider only those $\mathbf{h} \in \mathbb{C}^N$ such that

$$\mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\} \geq 0, \tag{3.10}$$

then our assumed scenario is an instantiation of the mean-shifted Gauss-Gauss problem [93, pp. 70]. In a mean-shifted Gauss-Gauss problem, the test statistic T is a real Gaussian random variable distributed according to

$$T \sim \begin{cases} \mathcal{N}\left(\mu_0, \sigma^2\right) & \text{under } \mathcal{H}_0 \\ \mathcal{N}\left(\mu_1, \sigma^2\right) & \text{under } \mathcal{H}_1 \end{cases}, \tag{3.11}$$

and the probability of false alarm (P_{FA}) can be kept constant by choosing the detection threshold according to

$$\gamma = Q^{-1}\left(P_{FA}\right)\sigma + \mu_0, \tag{3.12}$$

where Q is the complementary cumulative distribution function [93, pp. 21]. The probability of detection (P_D) is then given by

$$P_D = Q\left(Q^{-1}\left(P_{FA}\right) - \sqrt{d^2}\right), \tag{3.13}$$

where

$$d^2 := \frac{\left(\mu_1 - \mu_0\right)^2}{\sigma^2} \tag{3.14}$$

is known as the *deflection coefficient*. Note that P_D is a monotonically increasing function of d^2. Thus, the deflection coefficient is sufficient for characterizing detection performance, and should be made as large as possible insofar as it is under the system designer's control. For the problem considered in this section, the deflection coefficient is given by

$$d^2 = \frac{2\mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\}^2}{\mathbf{h}^H\mathbf{K}\mathbf{h}} \tag{3.15}$$

From this we observe that (3.10) is not restrictive because $\mathbf{h}$ and $-\mathbf{h}$ give rise to the same deflection coefficient for a given $\mathbf{s}$ and $\mathbf{K}$.

3.1.2 Signal-Filter Optimization

Of the parameters that define the deflection coefficient (3.15), only the receive filter ($\mathbf{h}$) and transmit signal ($\mathbf{s}$) are under a radar engineer's control. Making this dependence explicit, we can write the deflection coefficient as

$$d^2(\mathbf{s},\mathbf{h}) = (2\mathcal{E})\frac{\mathrm{Re}\left\{\widehat{\mathbf{h}}^H\widehat{\mathbf{s}}\right\}^2}{\widehat{\mathbf{h}}^H\mathbf{K}\widehat{\mathbf{h}}} \tag{3.16}$$

where $\widehat{\mathbf{s}}$ and $\widehat{\mathbf{h}}$ are unit vectors such that $\mathbf{s} = \sqrt{\mathcal{E}}\widehat{\mathbf{s}}$ and $\mathbf{h} = \sqrt{\mathcal{E}_h}\widehat{\mathbf{h}}$. Notice that d^2 does not depend upon the filter energy ($\mathcal{E}_h$), but it is monotonic in signal energy ($\mathcal{E}$). Thus, the maximum of d^2 with respect to $\widehat{\mathbf{s}}$ and $\widehat{\mathbf{h}}$ exists (i.e., the supremum is finite) only when $\mathcal{E}$ is restricted to be less than or equal to some finite value, say $\mathcal{E}_0$. This value is always known in practice, and the optimal signal will necessarily have energy equal to it.[1] Therefore, to optimize detection performance for a given $\mathbf{K}$, we need only find the unit vectors $\widehat{\mathbf{h}}$ and $\widehat{\mathbf{s}}$ that jointly maximize the deflection coefficient. The choice of $\mathcal{E}_0$ does not affect the optimal unit vectors. Thus, we can choose the convenient value of $\mathcal{E}_0 = 1/2$, and write the system

[1] Proof by contradiction: Let $\sqrt{\mathcal{E}_*}\widehat{\mathbf{s}}_*$ be the optimal waveform where $\mathcal{E}_* \in [0, \mathcal{E}_0)$. Then, by definition, $d^2(\mathcal{E}_*, \widehat{\mathbf{s}}; \mathbf{h}, \mathbf{K}) \geq d^2(\mathcal{E}_0, \widehat{\mathbf{s}}; \mathbf{h}, \mathbf{K})$. However, this implies $\mathcal{E}_* \geq \mathcal{E}_0$, which is a contradiction. □

design problem as

$$\underset{(\mathbf{s},\mathbf{h}) \in \mathcal{C}}{\arg\max} \ \frac{\mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\}^2}{\mathbf{h}^H\mathbf{K}\mathbf{h}} \tag{3.17}$$

where the constraint set $\mathcal{C}$ is defined by

$$\mathcal{C} := \left\{(\mathbf{s},\mathbf{h}) \in B_N \times \mathbb{C}^N : \mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\} \geq 0\right\} \tag{3.18}$$

Notice that we no longer use the "hat" notation because $\mathbf{s}$ is restricted to be in the unit ball (B_N), and d^2 does not depend on the filter energy. The solution to (3.17) is not necessarily unique. However, one solution is given by

$$(\mathbf{s},\mathbf{h}) = (\mathbf{u}_0,\mathbf{u}_0) \tag{3.19}$$

where $\mathbf{u}_0$ is the eigenvector of $\mathbf{K}$ corresponding to the minimum eigenvalue of $\mathbf{K}$.[2] We will refer to (3.19) as the *eigen-optimal* solution. In general, additional constraints will be placed on the transmit signal (e.g., constant modulus). Let $\mathcal{S} \subseteq B_N$ denote the set of signals that satisfy these additional constraints, then we can write the optimization problem as

$$\underset{(\mathbf{s},\mathbf{h}) \in \mathcal{D}}{\arg\max} \ \frac{\mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\}^2}{\mathbf{h}^H\mathbf{K}\mathbf{h}} \tag{3.20}$$

where

$$\mathcal{D} := \left\{(\mathbf{s},\mathbf{h}) \in \mathcal{S} \times \mathbb{C}^N : \mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\} \geq 0\right\}. \tag{3.21}$$

[2]This result is obtained by observing that $\mathbf{u}_0^H\mathbf{K}\mathbf{u}_0 \leq \mathbf{h}^H\mathbf{K}\mathbf{h}$ for all $\mathbf{h} \in B_N$, $\mathrm{Re}\left\{\mathbf{h}^H\mathbf{s}\right\}^2 = \left|\mathbf{h}^H\mathbf{s}\right|^2$ for all $(\mathbf{s},\mathbf{h}) \in \mathcal{C}$, and $\left|\mathbf{u}_0^H\mathbf{u}_0\right|^2 \geq \left|\mathbf{u}_0\mathbf{s}\right|^2$ for all $\mathbf{s} \in B_N$.

In practice, additional constraints may be placed on the signal/filter pair. This will be addressed in Sec. 3.2. For now, we assume (3.20) and (3.21) describe the signal/filter design problem.

3.1.3 Waveform-Only Optimization

Let $p(\mathbf{x}; \mathcal{H}_i)$ denote the pdf of a received signal $\mathbf{x}$ under hypothesis $\mathcal{H}_i$. Then the Neyman-Pearson theorem [93, pp. 64] says that in order to maximize P_D for a given P_{FA}, one should decide according to

$$L(\mathbf{x}) := \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{3.22}$$

where γ is found from

$$P_{FA} = \int_{\{\mathbf{x}: L(\mathbf{x}) > \gamma\}} p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} \tag{3.23}$$

It can be shown [93, pp. 478] that under the signal and interference assumptions made in the previous section, the Neyman-Pearson detector in (3.22) reduces to (3.6) with $\mathbf{h}$ replaced by

$$\mathbf{h}_{np} = \mathbf{K}^{-1}\mathbf{s}. \tag{3.24}$$

This is known as the *whitening filter*. Note that because $\mathbf{K}$ is positive definite, so too is $\mathbf{K}^{-1}$. This implies that $\mathrm{Re}\left\{\mathbf{h}_{np}^H \mathbf{s}\right\} = \mathbf{s}^H \mathbf{K}^{-1} \mathbf{s} > 0$ for any $\mathbf{s} \in \mathbb{C}^N$. As a result, $(\mathbf{s}, \mathbf{K}^{-1}\mathbf{s}) \in \mathcal{D}$ for all $\mathbf{s} \in \mathcal{S}$. Thus, if $(\mathbf{s}_\star, \mathbf{h}_\star)$ is a solution to (3.20), then so too is $(\mathbf{s}_\star, \mathbf{K}^{-1}\mathbf{s}_\star)$.[3] As such, we can simplify the optimization in (3.20) by solving only for the optimal waveform. The

[3] This is an immediate result of the Neyman-Pearson theorem. Another proof is to recognized that $\mathrm{Re}\left\{\mathbf{h}_{np}^H \mathbf{s}\right\} = \left|\mathbf{h}_{np}^H \mathbf{s}\right|$, and that for a given $\mathbf{s}_\star$, the deflection coefficient is a generalized Rayleigh quotient with $\mathbf{h}_{np}$ as the solution.

resulting problem formulation is given by

$$\arg\max_{\mathbf{s}\,\in\,\mathcal{S}} \ \mathbf{s}\mathbf{K}^{-1}\mathbf{s} \tag{3.25}$$

where the deflection coefficient corresponding to the Neyman-Pearson detector has been simplified to [93, pp. 479]

$$d^2_{np}(\mathbf{s}) := d^2(\mathbf{s}, \mathbf{K}^{-1}\mathbf{s}) = \mathbf{s}^H\mathbf{K}^{-1}\mathbf{s} \tag{3.26}$$

By specifying the form of the receive filter we have significantly reduced the dimension of the optimization problem. In what follows, we will also consider the matched filter case, where the filter is given by

$$\mathbf{h}_{mf} = \mathbf{s}. \tag{3.27}$$

The matched filter is the Neyman-Pearson filter when the interference is white (i.e., $\mathbf{K} = \mathbf{I}$). Again, we note that $\mathrm{Re}\left\{\mathbf{h}^H_{mf}\mathbf{s}\right\} = ||\mathbf{s}||^2 \geq 0$ for all $\mathbf{s} \in \mathbb{C}^N$. Thus, $(\mathbf{s}, \mathbf{s}) \in \mathcal{D}$ for all $\mathbf{s} \in \mathcal{S}$. The corresponding deflection coefficient is given by [93, pp. 476]

$$d^2_{mf}(\mathbf{s}) := d^2(\mathbf{s}; \mathbf{s}) = \frac{||\mathbf{s}||^4}{\mathbf{s}^H\mathbf{K}\mathbf{s}}. \tag{3.28}$$

If $\mathcal{S} = B_N$, or if $\mathcal{S}$ is such that $||\mathbf{s}||$ is constant for all $\mathbf{s} \in \mathcal{S}$, then the resulting optimization problem becomes

$$\arg\min_{\mathbf{s}\,\in\,\mathcal{S}} \ \mathbf{s}\mathbf{K}\mathbf{s}. \tag{3.29}$$

Notice that this is a minimization as opposed to the maximization in (3.25).

3.1.4 Waveform-Optimized Performance

If $(\mathbf{s}_0, \mathbf{h}_0)$ denotes a solution to a particular optimization problem, then $d^2(\mathbf{s}_0, \mathbf{h}_0)$ is called the *waveform-optimized performance* of the system. Let $(\mathbf{s}_j, \mathbf{h}_j)$ denote a solution to the joint signal/filter design problem (3.20), $\mathbf{s}_n$ denote a solution to whitening filter problem (3.25), and $\mathbf{s}_m$ denote a solution to the matched filter problem (3.29). Then, regardless of the constraint set $\mathcal{S}$, we have

$$d^2(\mathbf{s}_j, \mathbf{h}_j) = d^2_{np}(\mathbf{s}_n) \geq d^2_{mf}(\mathbf{s}_m). \tag{3.30}$$

That is, when the optimizations in (3.20), (3.25), and (3.29) are performed over the same constraint set ($\mathcal{S}$), then the waveform-optimized performance of the joint signal/filter design is no better than that of the whitening filter optimized design, and both designs are at least as good as the matched filter optimized design.

Note that $(\mathbf{s}_j, \mathbf{h}_j)$, $\mathbf{s}_n$, and $\mathbf{s}_m$ must be global optimizers if they are to be considered as solutions to their respective optimization problems. Global optimizers can be found in the literature for some scenarios. For example, [28] and [30] address a scenario involving the detection of a single non-moving point target at a known range and velocity in additive Gaussian interference and linear time invariant clutter. In that work, a closed-form solution to (3.25) is provided for a constraint set comprised of waveforms having a specified bandwidth and energy. However, the problem becomes non-convex when additional constraints are placed on the transmit signal (e.g., constant modulus), and only mathematical programming algorithms have been proposed for solving the resulting optimization problem. The output of such algorithms may be only locally optimal, and therefore, we have greater difficulty in predicting the waveform-optimized performance of each design. Furthermore, (3.30) may not hold when additional constraints are placed on the signal/filter pair (e.g., CAF constraints) because the constraint sets may be different for each problem. Thus, for highly constrained problems, the waveform-optimized performances of each architecture

must be evaluated and compared in order to determine which architecture provides the best performance. We present examples of this in the following sections.

3.2 Unknown Targets in Noise

In this section, we present a scenario in which the joint design, whitening filter design, and matched filter design must each be optimized over a different constraint set. In Sec. 3.3, we will solve for these problems and compare their respective waveform-optimized performances.

3.2.1 Problem Formulation

Consider the scenario in which a radar system must determine the number of non-moving point targets present in a scene.[4] Assume that the radar transmits a signal $s(t)$ with compact support over $[0, T]$ and negligible energy at frequencies $f \notin (f_c - B/2, f_c + B/2)$, where f_c and B are the center frequency and bandwidth, respectively. Assume this signal is reflected from each of the N_t point targets such that the combined return is given by

$$q(t) = \sum_{i=1}^{N_t} A_i s(t - \tau_i) \tag{3.31}$$

where A_i and τ_i are, respectively, the reflectivity coefficient and two-way propagation delay for target i. The received signal is given by

$$x(t) = q(t) + w(t) \tag{3.32}$$

where $w(t)$ is a WSS zero-mean additive Gaussian random process with known covariance matrix.

[4]The omission of Doppler effects does not change the nature of our examples, and as we will see, accounting for Doppler is conceptually straightforward, but computationally demanding. Thus, for simplicity, we assume the targets have no radial velocity with respect to the radar.

In contrast to the previous section, we assume that A_i and τ_i are unknown. The standard method of detecting targets in such a scenario is to collect the receive signal over some longer period ($T' > T$), convolve the received signal with a filter, and declare targets when the magnitude of the output exceeds a threshold. This correlation receiver is a common architecture for radars employing pulse compression [103, ch. 4][133], and it has a number of equivalent interpretations, such as 1) the generalized likelihood-ratio test (GLRT) for a single target with unknown amplitude and range [103, ch. 6][93], and 2) as an approximation to the inverse scattering problem involving multiple targets [134, ch. 4].

As before, we will assume that the receive signal is modulated to baseband, and that both the in-phase and quadrature signals are time sampled at a rate of B Hz. The discrete-time signal model is then $\mathbf{x} = \mathbf{q} + \mathbf{w}$, where $\mathbf{x}, \mathbf{q}, \mathbf{w} \in \mathbb{C}^M$. For simplicity, we assume that each target delay is an integer multiple of the sampling period so that

$$q_n = \sum_{i=1}^{Nt} A_i s_{n-k_i} \tag{3.33}$$

where $\mathbf{s} \in \mathbb{C}^N$ is the baseband discrete time transmit signal, and for target i, $A_i \in \mathbb{C}$ is the complex scattering coefficient, and k_i is the delay. Assuming the filter is the same length as the transmit signal, the cross-correlation (XCS) between the received signal ($\mathbf{x} \in \mathbb{C}^M$) and filter ($\mathbf{h} \in \mathbb{C}^N$) can be computed at each hypothesized range according to

$$R_k(\mathbf{x}, \mathbf{h}) = \sum_{n=0}^{N-1-k} x_{n+k} h_n^*. \tag{3.34}$$

The test for target presence at range k is then given by

$$|R_k(\mathbf{x}, \mathbf{h})| \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_k. \tag{3.35}$$

where the range-dependent threshold (γ_k) may be chosen by any of a number of criteria, such as cell-averaging CFAR [103, ch.7][135].

We now wish to formulate the signal/filter design problem so that system performance might be optimized. To do so, we must determine an appropriate objective function and constraint set. If we assume for a moment that there is only a single target present in the scene, then the post-filter signal-to-interference-plus-noise ratio (SINR) at the true target delay is given by

$$\text{SINR}\left(\mathbf{s},\mathbf{h}\right) = \frac{\left|\mathbf{h}^H\mathbf{s}\right|^2}{\text{E}\left\{\left|\mathbf{h}^H\mathbf{w}\right|\right\}^2} = \frac{\left|\mathbf{h}^H\mathbf{s}\right|^2}{\mathbf{h}^H\mathbf{K}\mathbf{h}} \tag{3.36}$$

Further, when either the Neyman-Pearson filter or the matched filter are used, then $\left|\mathbf{h}^H\mathbf{x}\right| = \text{Re}\left\{\mathbf{h}^H\mathbf{x}\right\}$ and $\text{SINR}\left(\mathbf{s},\mathbf{h}\right) = d^2(\mathbf{s},\mathbf{h})$. As before, it can be shown that the probability of detection is a monotonically increasing function of SINR [103, ch. 4]. Therefore, we choose SINR as our objective function.

When multiple targets are present in the return signal ($\mathbf{x}$), it is desirable for the cross-correlation between $\mathbf{s}$ and $\mathbf{h}$ to have a narrow main lobe so that closely spaced target returns can be resolved, and low side lobes so that weaker target returns are not masked by stronger returns. One way of imposing these constraints is to force the magnitude of the normalized XCS, given by

$$\tilde{R}_k(\mathbf{s},\mathbf{h}) = \frac{R_k(\mathbf{s},\mathbf{h})}{|\mathbf{h}^H\mathbf{s}|}, \tag{3.37}$$

to be below a masking sequence $\{m_k\}$ for $|k| = 1,\ldots,N-1$. Note that we do not constrain the XCS at $k = 0$ because $\tilde{R}_0(\mathbf{s},\mathbf{h})$ is always equal to unity. Attempting to constrain this value will unnecessarily slow the optimization process. Also note that the XCS is computed at negative lags by

$$\tilde{R}_{-k}(\mathbf{s},\mathbf{h}) = \tilde{R}_k^*(\mathbf{h},\mathbf{s}). \tag{3.38}$$

The masking sequence $\{m_k\}$ can be chosen so that the peak response is at $k = 0$, and so

that the XCS has a specific main lobe width and a specific peak side lobe ratio (PSLR). An example mask is presented in Sec. 3.3. Note that it is conceptually straightforward to extend this masking approach to constrain the signal/filter CAF.

In choosing the PSLR for the XCS constraint mask, it is helpful to consider the expected normalized XCS between the received signal and filter. This is given by

$$\mathrm{E}\left\{\left|\tilde{R}_k(\mathbf{x},\mathbf{h})\right|^2\right\} := \frac{\mathrm{E}\left\{\left|R_k(\mathbf{x},\mathbf{h})\right|^2\right\}}{\left|\mathbf{h}^H\mathbf{s}\right|^2} \tag{3.39}$$

$$= \frac{\left|R_k(\mathbf{q},\mathbf{h})\right|^2}{\left|\mathbf{h}^H\mathbf{s}\right|^2} + \frac{\mathbf{h}^H\mathbf{K}\mathbf{h}}{\left|\mathbf{h}^H\mathbf{s}\right|^2} \tag{3.40}$$

$$:= \left|\tilde{R}_k(\mathbf{q},\mathbf{h})\right|^2 + \frac{1}{\mathrm{SINR}\left(\mathbf{s},\mathbf{h}\right)} \tag{3.41}$$

The first term in (3.41) is due to the target returns, whereas the $1/\mathrm{SINR}\left(\mathbf{s},\mathbf{h}\right)$ term is due to the interference. We call the second term the XCS *noise floor*. If the PSLR is chosen too low, then the domain of optimization may be overly restricted, and the subsequent reduction in achievable SINR will manifest as an increase in the XCS noise floor. On the other hand, if the PSLR is chosen too high, then the achievable SINR will be increased, and the noise floor will reduce. However, weak target returns may then be masked by the side lobes of stronger returns.

In addition to the XCS constraints, we assume that the transmit signal must be constant modulus. This constraint can be explicitly addressed by designing only the phase function of the transmit signal. Hereafter, the vector $\mathbf{s}$ should be understood as a function of the phase vector $\boldsymbol{\phi}$, such that $s_n = a_n \exp\left(j\phi_n\right)$, where $a_n \in \mathbb{R}$ is the desired amplitude of the signal at time n. For constant modulus, $a_n = 1$ for all n. The constraint set for the joint signal/filter design can now be written as

$$\mathcal{D}_j := \left\{(\boldsymbol{\phi},\mathbf{h}) \in \mathbb{R}^N \times \mathbb{C}^N : \left|\tilde{R}_k(\mathbf{s},\mathbf{h})\right| \leq m(k),\ k \neq 0\right\}. \tag{3.42}$$

Notice that the dependence of the transmit vector $\mathbf{s}$ on the phase vector $\boldsymbol{\phi}$ is implicit. If we

assume a whitening filter architecture, then we can write the constraint set as

$$\mathcal{S}_n := \left\{ \boldsymbol{\phi} \in \mathbb{R}^N : \left| \tilde{R}_k(\mathbf{s}, \mathbf{K}^{-1}\mathbf{s}) \right| \leq m(k),\ k \neq 0 \right\}. \tag{3.43}$$

If a matched filter architecture is assumed, then the signal/filter XCS is symmetric about $k = 0$, and the constraint set becomes

$$\mathcal{S}_m := \left\{ \boldsymbol{\phi} \in \mathbb{R}^N : \left| \tilde{R}_k(\mathbf{s}, \mathbf{s}) \right| \leq m(k),\ k > 0 \right\}. \tag{3.44}$$

The resulting optimization problems can now be written as:

$$\underset{(\boldsymbol{\phi},\mathbf{h})\ \in\ \mathcal{D}_j}{\arg\max} \frac{\left|\mathbf{h}^H\mathbf{s}\right|^2}{\mathbf{h}^H\mathbf{K}\mathbf{h}^H} \tag{3.45}$$

$$\underset{\boldsymbol{\phi}\ \in\ \mathcal{S}_n}{\arg\max}\ \mathbf{s}^H\mathbf{K}^{-1}\mathbf{s} \tag{3.46}$$

$$\underset{\boldsymbol{\phi}\ \in\ \mathcal{S}_n}{\arg\min}\ \mathbf{s}^H\mathbf{K}\mathbf{s} \tag{3.47}$$

We will call a solution to (3.45) the *optimized joint design*, a solution to (3.46) the *whitening filter optimized* (WF-optimized) design, and a solution to (3.47) the *matched filter optimized* (MF-optimized) design.

3.2.2 Remarks

Before proceeding to numerical examples, let us consider a few observations regarding the optimization problems in (3.45)-(3.47). First, notice that all three problems can be thought of as joint signal/filter optimizations, where the form of the filter is restricted in (3.46) and (3.47), but not in (3.45). Second, the domain of optimization is different in each problem. This is caused by the constraints imposed on the interaction between $\mathbf{s}$ and $\mathbf{h}$. If we let

$(\mathbf{s}_j, \mathbf{h}_j)$ be an optimized joint design, and $\mathbf{s}_n$ be a WF-optimized design, then clearly

$$\text{SINR}\left(\mathbf{s}_j, \mathbf{h}_j\right) \geq \text{SINR}\left(\mathbf{s}_n, \mathbf{K}^{-1}\mathbf{s}_n\right). \tag{3.48}$$

However, this result does not necessarily mean joint signal/filter design is preferable in all circumstances. This is because problems (3.45)-(3.47) must be solved numerically, and therefore, the choice of algorithm and initialization play an important role.

Problems (3.45)-(3.47) can each be solved using standard nonlinear programming algorithms such as sequential quadratic programming (SQP) or interior point methods [131]. However, the results may not be global optimizers. Further, SQP and interior point methods can fail if they are not initialized with a feasible (i.e., interior) point. As we will show in the next section, it will usually be easier to initialize (3.45) and (3.47) with a feasible point than it is to initialize (3.46) properly. Examining constraint sets (3.42)-(3.44) we see that the joint signal/filter design has $3N$ design variables (signal phase plus the real and imaginary parts of the filter), whereas (3.46) and (3.47) each have only N design variables . Furthermore, (3.47) has only $N-1$ constraints, whereas (3.45) and (3.46) have $2(N-1)$. Finally, computing the the cost and constraint gradients is more expensive for (3.45) than for (3.46) and (3.47). (See Appendix.) From these observations it appears that the MF-optimized design is the most computationally tractable.

3.3 Examples

3.3.1 Overview

In this section we present solutions to problems (3.45)-(3.47) for two example interference processes. For each example, the signal and filter were of length $N = 64$, and the signal was constant modulus. Each problem was solved using the MATLAB Optimization Toolbox [136]. In order to provide a fair comparison, each problem was solved using both the

SQP algorithm [136, pp. 4-26] and the interior point algorithm [136, pp. 4-36]. Only the results for the most efficient algorithm are reported. The algorithms were allowed to perform 5000 major iterations before termination. Although, they were allowed to terminate early if a local optimum was found, or a feasible design could not be achieved. Problem (3.45) and (3.46) were solved most efficiently by the interior point algorithm, whereas (3.47) was solved most efficiently by SQP. An initial barrier parameter of 0.1 was used for the interior point algorithm. Note that all of the gradients and Jacobians required by these algorithms are provided in the Appendix.

In order to assess the waveform-optimized performance of the three designs, we compare their SINR to that of three baseline signal/filter pairs. The first baseline signal/filter pair, which we shall refer to as the "LFM/MF" design, is a linearly frequency modulated (LFM) waveform ($\mathbf{s}_L$) with 70% bandwidth, and the associated matched filter. The measure of performance for each design is the amount of SINR provided in excess of the LFM/MF SINR. This gain is calculated as

$$G(\mathbf{s}, \mathbf{h}) := \frac{\text{SINR}\,(\mathbf{s}, \mathbf{h})}{\text{SINR}\,(\mathbf{s}_L, \mathbf{s}_L)} \tag{3.49}$$

We note that both the SQP and interior-point algorithms require initial points that are at least close to feasible. In the examples that follow, the normalized XCS mask was chosen so that the LFM/MF design was feasible. Problem (3.45) was initialized with $(\mathbf{s}_L, \mathbf{s}_L)$, whereas problems (3.46) and (3.47) were initialized with $\mathbf{s}_L$. Thus, the optimized joint design and the MF-optimized design were initialized with feasible points, whereas the WF-optimized design was not. A related feasibility problem could be solved in order to initialize (3.46) with a point in $\mathcal{S}_n$, but this would only add to the time required to solve for the WF-optimized design.

The second baseline signal/filter pair that we consider consists of the LFM signal and the corresponding whitening filter. We call this the "LFM/WF" design. The third signal/fil-

ter pair is the eigen-optimal design given in (3.19). Together, the second and third baseline solutions provide insight into the respective contributions of filtering and signal design to the optimization process. When the interference has a spectrum that is dissimilar to the baseline LFM, then it can be effectively filtered, and we expect the LFM/WF approach to perform nearly as well as the eigen-optimal solution. For the same reason, we expect the WF-optimized solution to perform well compared to the other solutions. When the interference is similar to the baseline LFM, then it cannot be easily filtered. In this case, the eigen-optimal gain will be much higher than the LFM/WF gain, and one expects waveform optimization to play a more significant role in performance improvement.

3.3.2 Dissimilar Interference

We now consider an example in which the interference is dissimilar to the baseline LFM. Assume that the interference consists of a white Gaussian random process (i.e., the noise) added to an autoregressive random process (i.e., the interference) whose transfer function is given by

$$H(z) = \frac{1}{(1 - 1.5z^{-1} + 0.7z^{-2})^4}. \tag{3.50}$$

Assume the two processes are scaled so that the interference-to-noise ratio (INR) is 40 dB and the SINR is -15 dB. Hereafter, we shall refer to the noise-plus-interference as simply the interference. The interference covariance matrix $\mathbf{K}$ can be found in closed form [137], and we assume it is known.

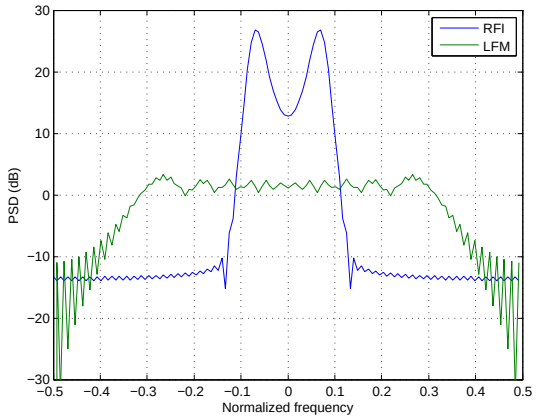
The interference power spectral density is shown in Fig. 3.1 along with the PSDs of the three baseline designs. Note that the receive filters in Figs. 3.1(b) and 3.1(c) have been scaled so that $\left|\mathbf{h}^H\mathbf{s}\right| = 1$. Allowing for this scaling, one can observe that the whitening filter shown in Fig. 3.1(b) is matched to the LFM at those frequencies at which the LFM PSD is greater than the interference PSD. This is a well known result that can be explained

by the fact that autocorrelation matrices can be diagonalized by the Fourier basis [138]. This also explains why the eigen-optimal solution shown in Fig. 3.1(c) places all of the transmit energy into the low-noise portion of the band (corresponding to the low-noise subspace of $\mathbf{K}$).

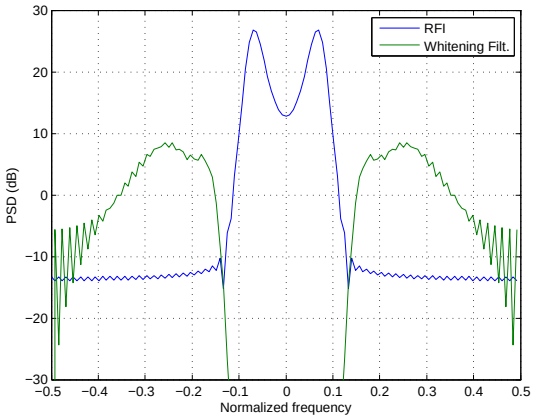
As expected, filtering alone provides a great deal of gain for this example. The LFM/WF solution provided 38.8 dB of gain, whereas the eigen-optimal solution provided only slightly more at 41.5 dB. However, as shown in Fig. 3.2, neither of these baseline solutions has a good XCS. Let us assume that the -20 db PSLR of the LFM/MF solution is acceptable. How should the XCS mask be chosen for the optimized designs? To answer that question, we note that the noise floors of the expected LFM/WF XCS and the expected eigen-optimal XCS are, respectively, 38.8 dB and 41.5 dB below the noise floor of the expected LFM/MF XCS. Further, the LFM/WF design is nearly feasible, with only the near-in sidelobes being greater than -20 dB. As such, we do not expect a PSLR of -20 dB to be overly restrictive.

The XCS mask chosen for this example is shown in Fig. 3.3 along with the final signal/-filter XCS for each optimized design. A feasible design was found for each problem, even for (3.46), which was initialized with the infeasible design shown in 3.2(b). Each solution provided at least 35 dB of gain above the LFM/MF. This translates into an expected XCS noise floor of at least 35 dB below the expected LFM/MF noise floor. Thus, the -20 dB PSLR was a prudent choice for this example.

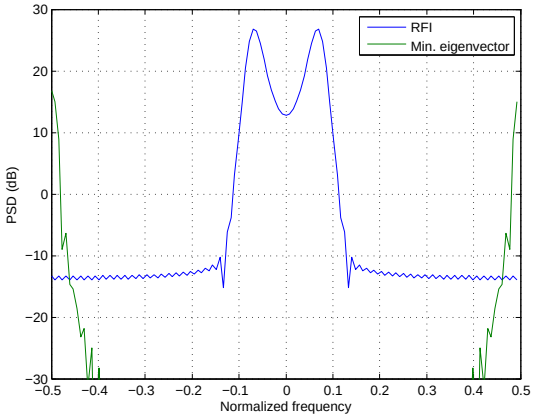
The power spectral densities of the optimized solutions are shown in Fig. 3.4. Again, the filters have been scaled so that $\left|\mathbf{h}^H\mathbf{s}\right| = 1$. As expected, the MF-optimized design places its energy outside of the strong interference region, providing 38.4 dB of gain. The WF-optimized design and the optimized joint signal/filter design have similar behavior, with each design achieving the maximum gain of 41.3 dB. The performance improvement over time for each optimized design is shown in Fig. 3.5. From this figure we can see that the WF-optimized design not only provided maximum gain, but it did so in far less time than the other solutions.



(a) LFM

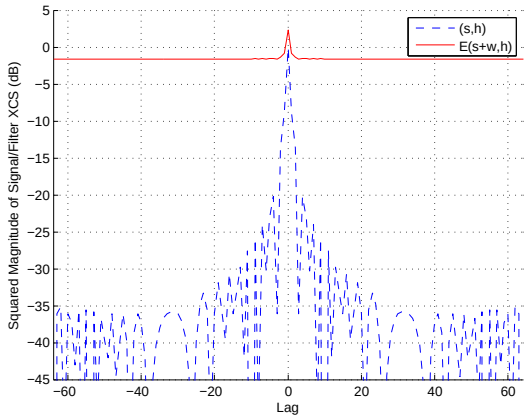


(b) Whitening filter.

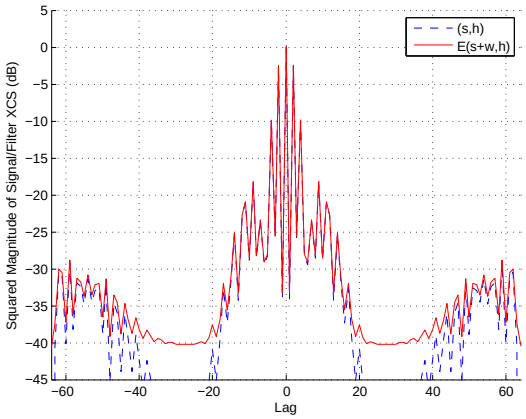


(c) Minimum eigenvector.

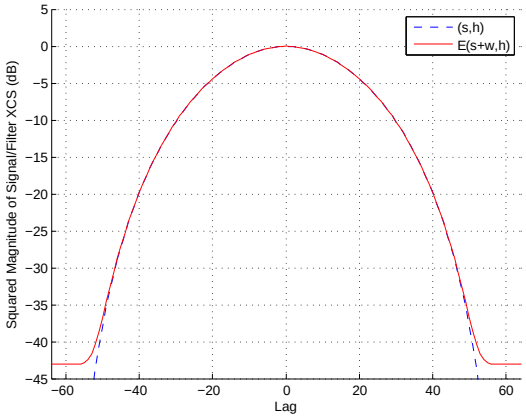
Figure 3.1: The power spectral density of the dissimilar interference is shown along with the PSD of (a) the baseline LFM, (b) the corresponding whitening filter, and (c) the minimum eigenvector. For this case, the LFM/WF design provides 38.8 dB of gain above the LFM/MF, whereas the eigen-optimal design provides 41.5 dB of gain.



(a) LFM/MF.

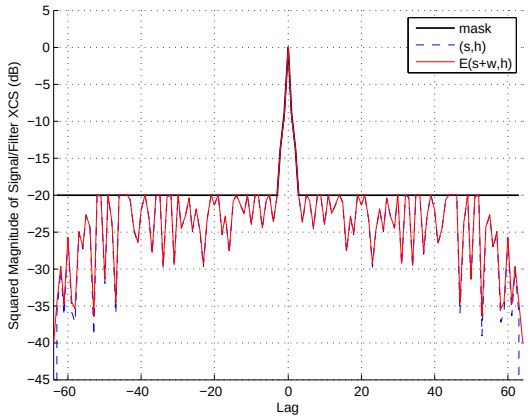


(b) LFM/WF.

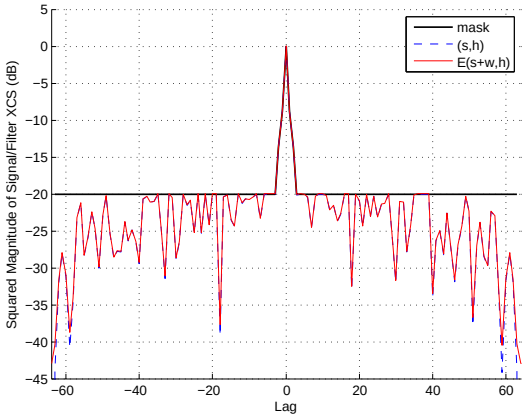


(c) Eigen-optimal signal/filter.

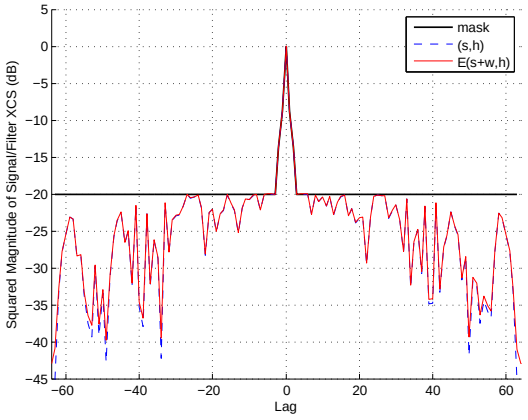
Figure 3.2: The squared magnitude of the normalized signal/filter XCS for each baseline design in the dissimilar interference case: (a) LFM/MF, (b) LFM/WF, (c) eigen-optimal signal/filter. In each subfigure, (s,h) denotes the XCS between the transmit signal and filter, and E(s+w,h) denotes the expected XCS between the received signal and filter. For this case, the LFM/WF design provides 38.8 dB of gain above the LFM/MF, whereas the eigen-optimal design provides 41.5 dB of gain.



(a) MF-optimized design.

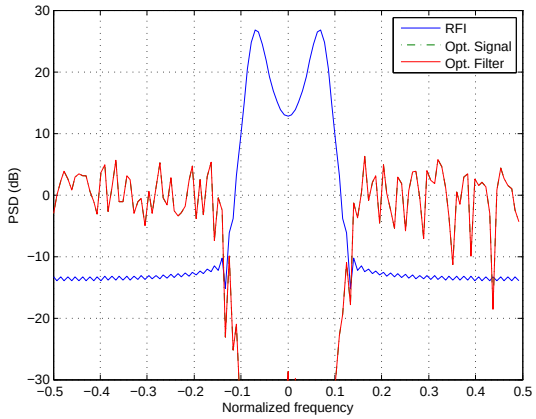


(b) WF-optimized design.

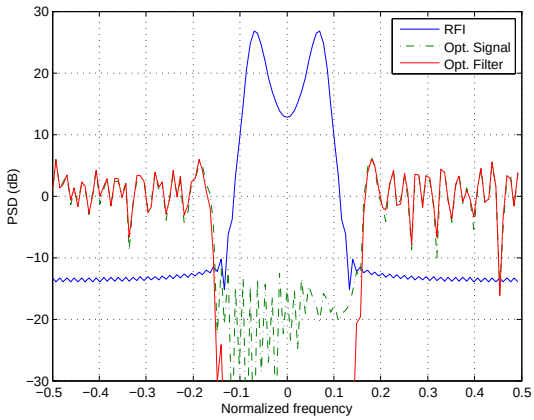


(c) Optimized joint signal/filter design.

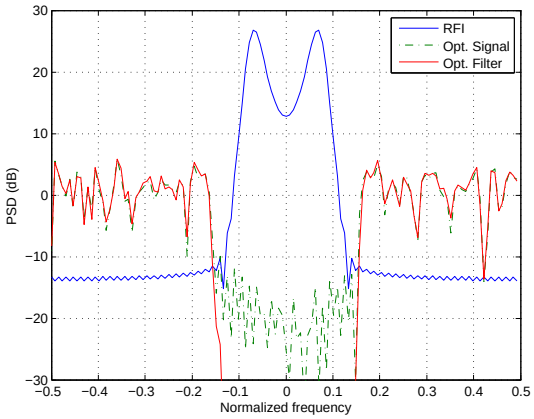
Figure 3.3: The -20 dB PSLR XCS mask is shown along with the squared magnitude of the normalized signal/filter XCS for each optimized solution in the dissimilar interference case: (a) the MF-optimized design, (b) the WF-optimized design, and (c) the optimized joint signal/filter design. In each subfigure, (s,h) denotes the XCS between the transmit signal and filter, and E(s+w,h) denotes the expected XCS between the received signal and filter. For this case, the MF-opt. solution provides 38.4 dB of gain above the LFM/MF, the WF-opt. design provides 41.3 dB, and the opt. joint signal/filter design provides 41.3 dB of gain.



(a) MF-optimized design.

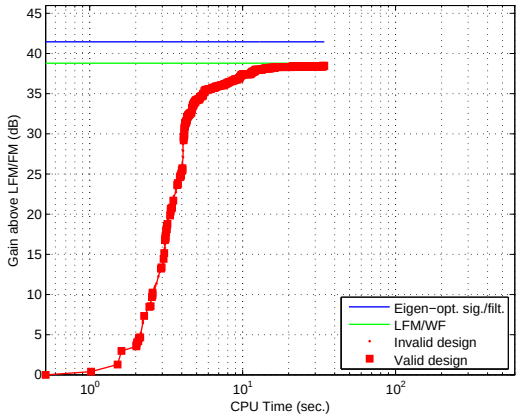


(b) WF-optimized design.

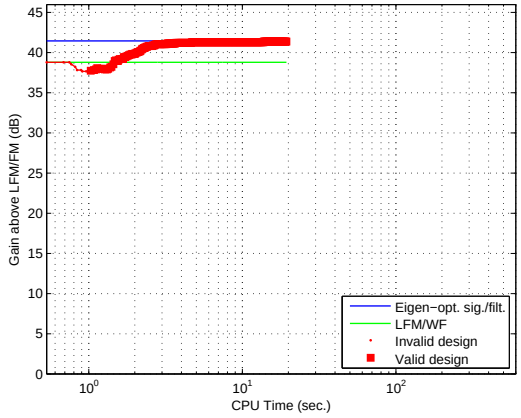


(c) Optimized joint signal/filter design.

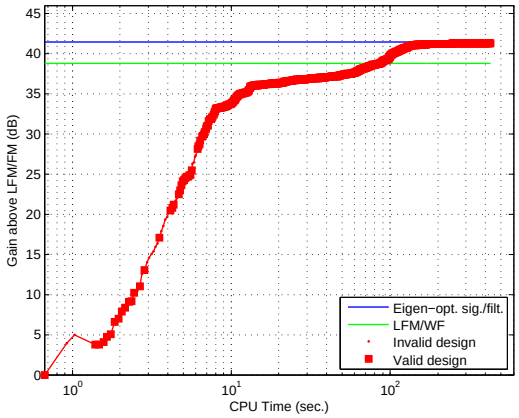
Figure 3.4: The power spectral density of the dissimilar interference is shown along with the PSD of (a) the MF-optimized design, (b) the WF-optimized design, and (c) the optimized joint signal/filter design. For this case, the MF-opt. solution provides 38.4 dB of gain above the LFM/MF, the WF-opt. design provides 41.3 dB, and the opt. joint signal/filter design provides 41.3 dB of gain.



(a) MF-optimized design.



(b) WF-optimized design.



(c) Optimized joint signal/filter design.

Figure 3.5: Performance improvement over time for the dissimilar interference case. Each subfigure shows the gain at each major iteration for (a) the MF-optimized design, (b) the WF-optimized design, and (c) the optimized joint signal/filter design. The lower solid line denotes the gain provided by the LFM/WF solution (38.8 dB), and the upper solid line denotes the gain provided by the eigen-optimal solution (41.5 dB). Optimized designs that violate the constraints (i.e., invalid designs) are denoted by dots, whereas designs that do not violate the constraints (i.e., valid designs) are denoted by squares.

3.3.3 Similar Interference

We now consider an example in which the interference is similar to the baseline LFM. Assume that the interference consists of a white Gaussian random process added to another white Gaussian random process that has been band-pass filtered. As before, the two processes are scaled to provide a 40 dB INR and a -15 dB SINR. For this example, the interference covariance matrix ($\mathbf{K}$) was estimated using one million realizations. The interference power spectral density is shown in Fig. 3.6 along with the PSDs of the three baseline solutions. Note that the filters in Figs. 3.6(b) and 3.6(c) have been scaled so that $\left|\mathbf{h}^H\mathbf{s}\right| = 1$. As expected, we observe that the whitening filter is matched to the LFM at those frequencies at which the LFM PSD is greater than the interference PSD. However, because the interference is so similar to the baseline LFM, the whitening filter must reject most of the signal energy. Therefore, while the deflection coefficient denominator $\mathbf{h}^H\mathbf{K}\mathbf{h}$ is made small, the numerator $\mathbf{h}^H\mathbf{s}$ is also small. This explains the large filter magnitude in the low-noise portion of the band. For this example, the LFM/WF solution provided 11.5 dB of gain above the LFM/MF solution. However, the eigen-optimal solution provided 41.7 dB of gain. Thus, we expect waveform optimization to play a significant role in performance improvement.

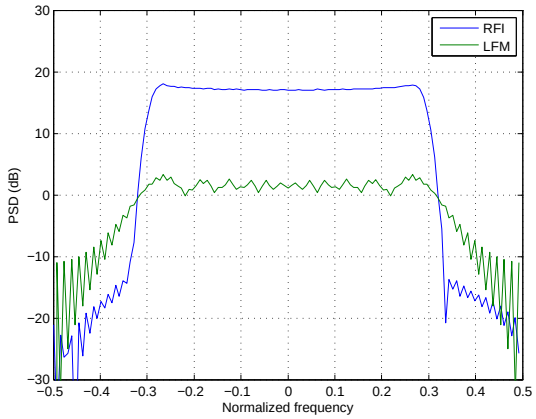
As before, the LFM/WF and eigen-optimal solutions have poor XCSs. (Fig. 3.7.) If we assume that the -20 db PSLR of the LFM/MF solution is acceptable, how should the XCS mask be chosen? The eigen-optimal noise floor (-41.7 dB) makes it conceivable that an optimized solution could achieve a -20 dB noise floor. However, the LFM/WF noise floor (-11.5 dB) is much higher, and the LFM/WF sidelobes are all greater than -20 dB. In this case, it is not clear how the max PSLR should be chosen. So, we choose -20 dB.

The XCS mask for this example is shown in Fig. 3.8 along with the final signal/filter XCS for each optimized solution. As can be seen, a feasible solution was found for the MF-optimized and optimized joint signal/filter designs, but not for the WF-optimized design. Of the two feasible designs, the MF-optimized design provided only 5.2 dB of gain,

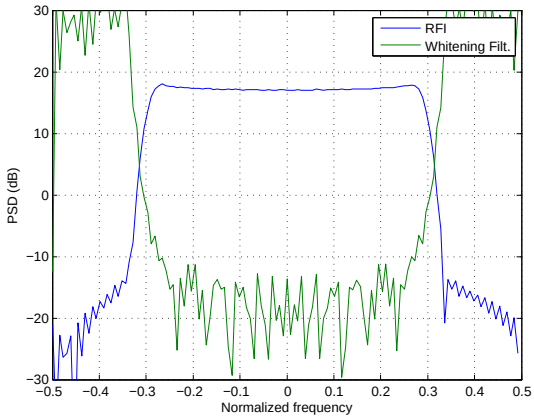
whereas the optimized joint design provided 10.0 dB. The power spectral densities of the optimized solutions are shown in Fig. 3.9. Notice that each design must place a significant amount of energy in the interference portion of the band in order to satisfy the constraints. This explains why the resulting gains are lower than in the previous example. The performance improvement over time for each optimized design is shown in Fig. 3.10. As in the previous example, the joint design appears to be as computationally efficient as the MF-optimized approach. Also, we can see that the WF-optimized problem terminated without ever providing a feasible design. This is a characteristic difficulty with the WF-optimized approach: determining a feasible solution for initialization. Of course, one can attempt to solve the relevant feasibility problem to find a suitable initial design, but this would only add to the computational complexity of the solution.

3.4 Conclusion

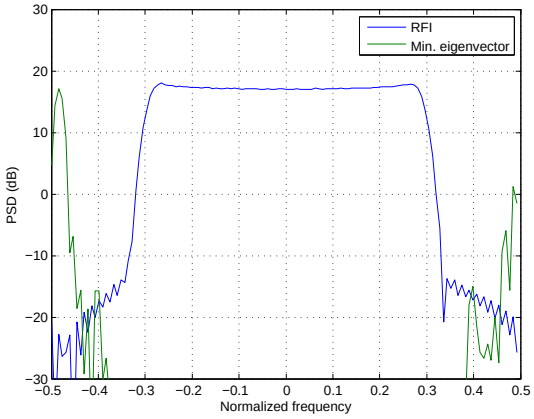
The examples in the previous section demonstrate that the whitening filter is not always the best choice when the transmit waveform is optimized for a particular interference environment. In the dissimilar interference example, the WF-optimized solution achieved the upper bound on gain in a minimal amount of time, despite the fact that the algorithm used to solve the optimization problem was initialized with an infeasible design. However, in the similar interference example, a feasible solution was never found for the whitening filter architecture. This difficulty with initialization is a distinct disadvantage of the whitening filter approach. On the other hand, the extra degrees of freedom available in a joint signal/filter design make it easier for the optimizing algorithm to satisfy the constraints. Thus, a joint design is more likely to provide a superior solution. In both examples, the joint signal/filter design provided more gain than the MF-optimized design, and it did so at a comparable computational cost. For these reasons, a joint signal/filter design may be preferable in many applications.



(a) LFM

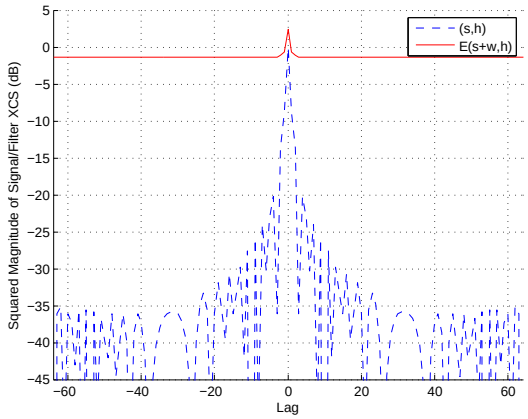


(b) Whitening filter.

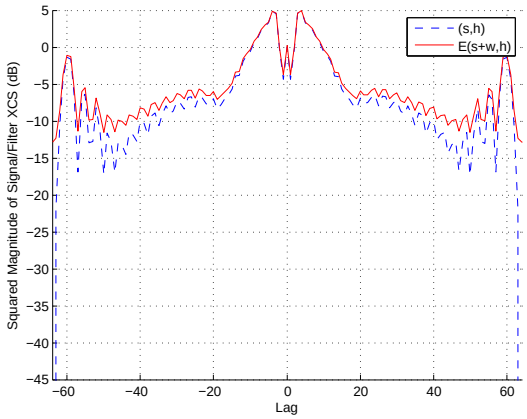


(c) Minimum eigenvector.

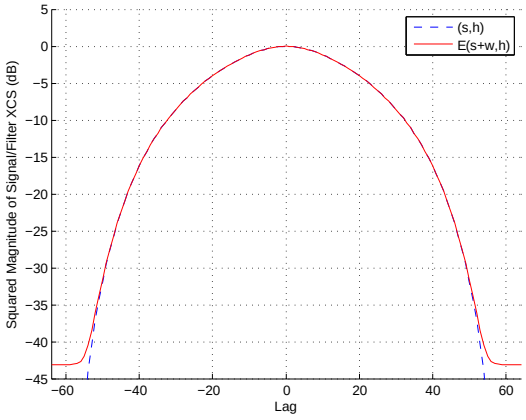
Figure 3.6: The power spectral density of the similar interference is shown along with the PSD of (a) the baseline LFM, (b) the corresponding whitening filter, and (c) the minimum eigenvector. For this case, the LFM/WF design provides 11.5 dB of gain above the LFM/MF, whereas the eigen-optimal design provides 41.7 dB of gain.



(a) LFM/MF.

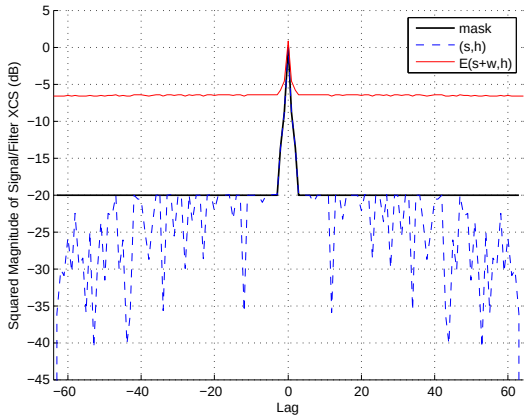


(b) LFM/WF.

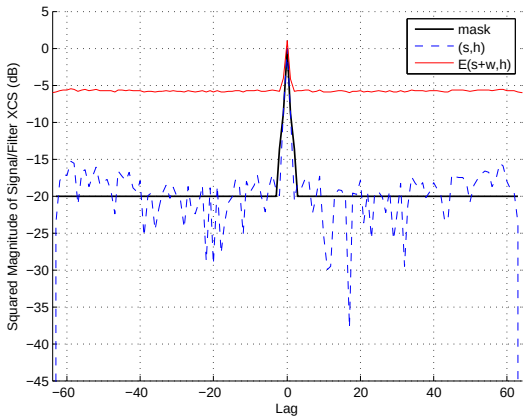


(c) Eigen-optimal signal/filter.

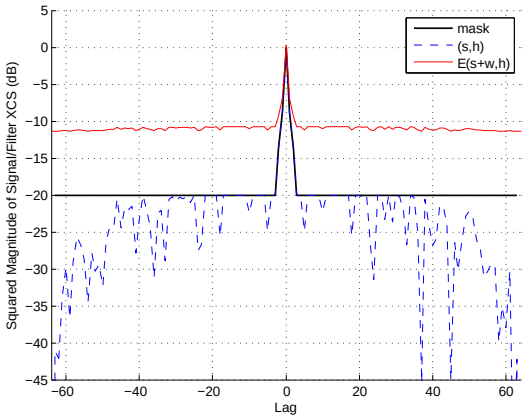
Figure 3.7: The squared magnitude of the normalized signal/filter XCS for each baseline design in the similar interference case: (a) LFM/MF, (b) LFM/WF, (c) eigen-optimal signal/filter. In each subfigure, (s,h) denotes the XCS between the transmit signal and filter, and E(s+w,h) denotes the expected XCS between the received signal and filter. For this case, the LFM/WF design provides 11.5 dB of gain above the LFM/MF, whereas the eigen-optimal design provides 41.7 dB of gain.



(a) MF-optimized design.

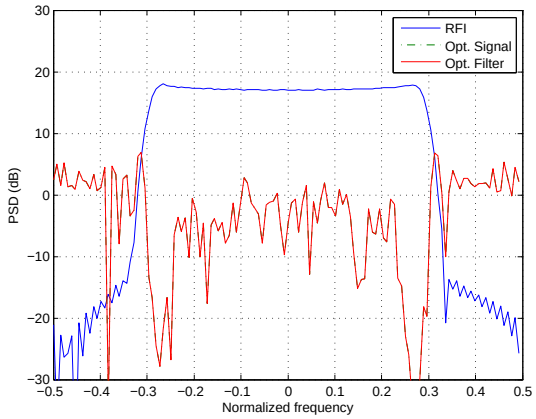


(b) WF-optimized design.

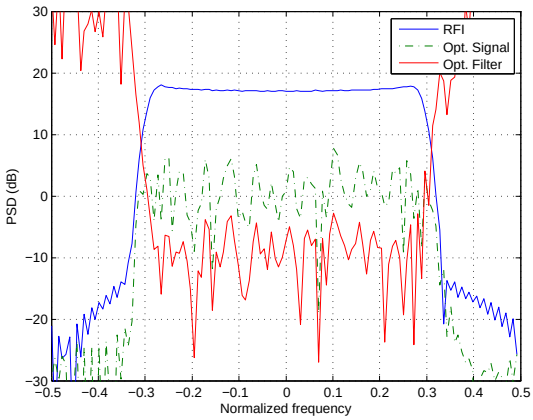


(c) Optimized joint signal/filter design.

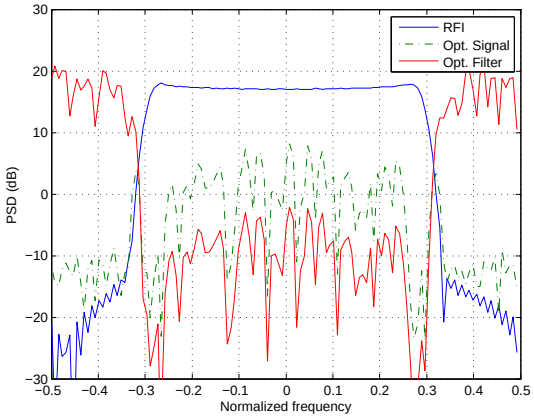
Figure 3.8: The -20 dB PSLR XCS mask is shown along with the squared magnitude of the normalized signal/filter XCS for each optimized solution in the similar interference case: (a) the MF-optimized design, (b) the WF-optimized design, and (c) the optimized joint signal/filter design. In each subfigure, (s,h) denotes the XCS between the transmit signal and filter, and E(s+w,h) denotes the expected XCS between the received signal and filter. For this case, the MF-opt. solution provides 5.2 dB of gain above the LFM/MF, the WF-opt. design provides 4.6 dB, and the opt. joint signal/filter design provides 10.0 dB of gain.



(a) MF-optimized design.

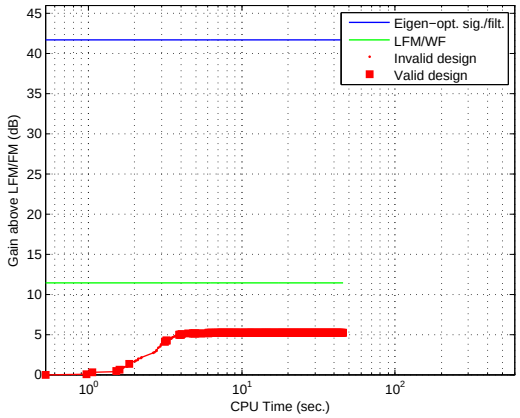


(b) WF-optimized design.

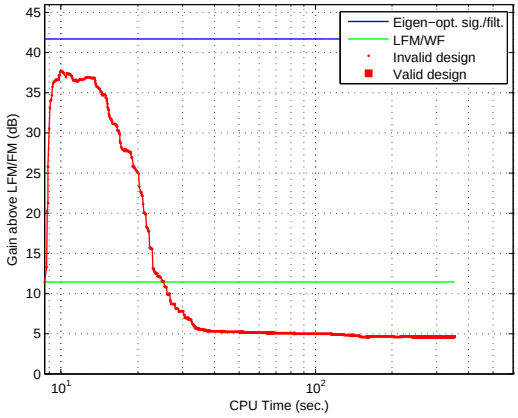


(c) Optimized joint signal/filter design.

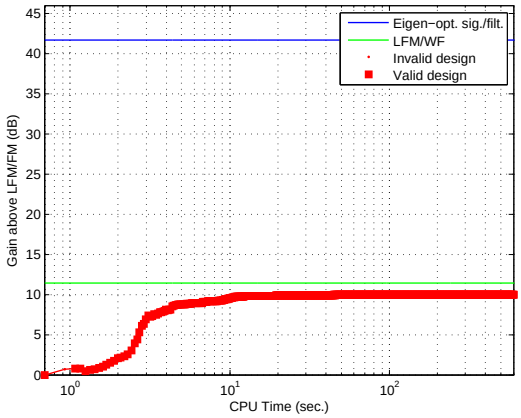
Figure 3.9: The power spectral density of the similar interference is shown along with the PSD of (a) the MF-optimized design, (b) the WF-optimized design, and (c) the optimized joint signal/filter design. For this case, the MF-opt. solution provides 5.2 dB of gain above the LFM/MF, the WF-opt. design provides 4.6 dB, and the opt. joint signal/filter design provides 10.0 dB of gain.



(a) MF-optimized design.



(b) WF-optimized design.



(c) Optimized joint signal/filter design.

Figure 3.10: Performance improvement over time for the similar interference case. Each subfigure shows the gain at each major iteration for (a) the MF-optimized design, (b) the WF-optimized design, and (c) the optimized joint signal/filter design. The lower solid line denotes the gain provided by the LFM/WF solution (11.5 dB), and the upper solid line denotes the gain provided by the eigen-optimal solution (41.7 dB). Optimized designs that violate the constraints (i.e., invalid designs) are denoted by dots, whereas designs that do not violate the constraints (i.e., valid designs) are denoted by squares.

Chapter 4

Computationally Efficient Formulations

In the previous chapter, we introduced a direct approach to constraining the signal/filter XCS. However, the algorithms used to solve the resulting problem formulations might be be too computationally expensive for some applications. This might occur when the interference is not strictly WSS, but it can be approximated as such over a short interval. As such, we seek suboptimal solutions that are less computationally intensive. To the best of our knowledge, there are no other approaches for constraining the signal/filter XCS other than the direct approach. However, if we assume a matched filter receiver, then we must constrain only the autocorrelation sequence (ACS) of the transmit signal, and indirect methods for constraining a waveform's ACS do exist. One such indirect method uses a *similarity constraint*, which forces the solution to be close in the Euclidean sense to some other waveform that possesses a desirable ACS [33, 35]. This approach can be motivated by Sussman's work [139], which shows that the error (Euclidean distance) between the ACSs of two waveforms decreases as the error between the two waveforms decreases. The utility of the similarity constraint is explored in Sec. 4.1. In Sec. 4.2, we discuss another indirect approach based on a parametrization of nonlinear frequency modulation. We note that while only the waveform's ACS is addressed explicitly, both approaches can be used to constrain the waveform's ambiguity function.

4.1 ACS and Modulus Constraints

In this section, we investigate the efficacy of the similarity constraint when waveform modulus is also constrained. In Sec. 4.1.1 and 4.1.2, we describe two methods of constraining the waveform ACS and three methods of constraining the waveform modulus. This results in six possible nonlinear programming problems. Standard numerical algorithms are then used to solve each formulation for an example interference process, and the results are presented in Sec. 4.1.3. Concluding remarks for this section are given in Sec. 4.1.4.

4.1.1 Magnitude and Phase Design

In this section, we consider the design of both the real and imaginary parts of a complex-valued waveform. This will allow for amplitude and phase modulation, as opposed to the phase-only modulation described in the previous chapter. Since we assume a matched filter architecture, the post-filter SINR is given by the functional $F : \mathbb{C}^N \rightarrow \mathbb{R}$ defined by

$$F(\mathbf{s}) := \frac{\left|\mathbf{s}^H\mathbf{s}\right|^2}{\mathbf{s}^H\mathbf{K}\mathbf{s}}, \tag{4.1}$$

where $\mathbf{s} \in \mathbb{C}^N$ is the waveform to be optimized and $\mathbf{K} \in \mathbb{C}^{N\times N}$ is a positive definite matrix. Let $\tilde{R}_k : \mathbb{C}^N \rightarrow \mathbb{R}$ map a waveform to the k^{th} lag of its normalized ACS. That is,

$$\tilde{R}_k(\mathbf{s}) := \frac{1}{|\mathbf{s}^H\mathbf{s}|} \sum_{n=0}^{N-1-k} s_{n+k}s_n^*. \tag{4.2}$$

As in Ch. 3, we can constrain the peak sidelobe height and mainlobe width of the normalized ACS by forcing the intensity of $\tilde{R}_k(\mathbf{s})$ to be less than some value $m_k \in \mathbb{R}_+$ at each lag k. To do so, we define the k^{th} ACS constraint functional $A_k : \mathbb{C}^N \rightarrow \mathbb{R}$ to be

$$A_k(\mathbf{s}) := \left|\tilde{R}_k(\mathbf{s})\right|^2 - m_k, \tag{4.3}$$

and we define a combined ACS constraint function $\mathcal{A} : \mathbb{C}^N \rightarrow \mathbb{R}^N$ by

$$\mathcal{A}(\mathbf{s}) := [A_1(\mathbf{s}) \ \ldots \ A_{N-1}(\mathbf{s})]^T. \tag{4.4}$$

Subjecting the maximization of (4.1) to the condition $\mathcal{A}(\mathbf{s}) \preceq \mathbf{0}$, where $\mathbf{0}$ is the zero vector, will directly impose constraints on the waveform ACS. Note that we do not constrain the ACS at $k = 0$. This is because $\tilde{R}_0(\mathbf{s})$ is always equal to unity, and constraining this value will unnecessarily slow the optimization process.

As an alternative to the ACS mask approach, we can indirectly impose ACS constraints on the maximization of (4.1) by using the similarity constraint functional $\mathcal{S} : \mathbb{C}^N \rightarrow \mathbb{R}$ defined by

$$\mathcal{S}(\mathbf{s}) := ||\mathbf{s}_d - \mathbf{s}||^2 - \epsilon \tag{4.5}$$

where $\mathbf{s}_d \in \mathbb{C}^N$ is a vector with a desirable ACS, and $\epsilon > 0$ controls the Euclidean distance of $\mathbf{s}$ from $\mathbf{s}_d$. Replacing the condition $\mathcal{A}(\mathbf{s}) \preceq \mathbf{0}$ with $\mathcal{S}(\mathbf{s}) \leq 0$ will constrain the ACS of the minimizer of (4.1) to be close to the ACS of $\mathbf{s}_d$. Just how close depends on the similarity constant ϵ. There is currently no formula for precisely relating ϵ to the distance between autocorrelation sequences. Therefore, we must chose ϵ based on experimentation.[1]

It was shown in Ch. 2 that the peak power of the waveform must be constrained even when the waveform is not restricted to be constant modulus. To control the maximum waveform modulus, we define the n^{th} waveform modulus constraint functional $M_n : \mathbb{C}^N \rightarrow \mathbb{R}$ by

$$M_n(\mathbf{s}) := |s_n|^2 - p_n, \tag{4.6}$$

[1]Sussman's work relating the Euclidean distance between ambiguity functions in $L^2(\mathbb{R}^2)$ to the Euclidean distance between signals in $L^2(\mathbb{R})$ might explain how the similarity constraint indirectly constrains the waveform ACS when waveform energy is held constant [139]. However, an analytical method of determining the similarity constant has not been reported.

where $p_n \in \mathbb{R}_+$ represents the maximum allowable squared magnitude (i.e., peak power) at time n. We then combine all modulus constraints into a single function $\mathcal{M} : \mathbb{C}^N \to \mathbb{R}^N$ defined by

$$\mathcal{M}(\mathbf{s}) := [M_0(\mathbf{s}) \ \ldots \ M_{N-1}(\mathbf{s})]^T. \tag{4.7}$$

The maximum modulus of the waveform can then be constrained by subjecting the maximization of (4.1) to the condition $\mathcal{M}(\mathbf{s}) \preceq \mathbf{0}$.

4.1.2 Phase-Only Design

If the waveform must be constant modulus, or the waveform must have some desired taper, we can choose each p_n appropriately, and replace $\mathcal{M}(\mathbf{s}) \preceq \mathbf{0}$ with $\mathcal{M}(\mathbf{s}) = \mathbf{0}$. In this case, we need only find the optimal phase vector $\boldsymbol{\phi} \in \mathbb{R}^N$ such that

$$s_n = \sqrt{p_n}\exp\left(j\phi_n\right) \tag{4.8}$$

The numerator in (4.1) does not vary with $\boldsymbol{\phi}$. Thus, maximizing (4.1) is equivalent to *minimizing* $\widehat{F} : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\widehat{F}(\boldsymbol{\phi}) := \mathbf{s}^H\mathbf{K}\mathbf{s}, \tag{4.9}$$

where the dependence of $\mathbf{s}$ on $\boldsymbol{\phi}$ is implicit. Similarly, the denominator in (4.2) does not vary with $\boldsymbol{\phi}$. Therefore, by scaling each p_n and m_k appropriately, we can use the unscaled ACS of $\mathbf{s}$, which is given by

$$R_k(\mathbf{s}) := \sum_{n=0}^{N-1-k} s_{n+k}s_n^*. \tag{4.10}$$

We can then define the k^{th} ACS constraint functional $\widehat{A}_k : \mathbb{R}^N \rightarrow \mathbb{R}$ by

$$\widehat{A}_k(\boldsymbol{\phi}) := |R_k(\mathbf{s})|^2 - m_k, \tag{4.11}$$

and the combined constraint function $\widehat{\mathcal{A}} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ by

$$\widehat{\mathcal{A}}(\boldsymbol{\phi}) := [\widehat{A}_1(\boldsymbol{\phi}) \ \ldots \ \widehat{A}_{N-1}(\boldsymbol{\phi})]^T. \tag{4.12}$$

Together, (4.11) and (4.12) can be used to directly constrain the waveform ACS. The similarity constraint can be imposed by using

$$\widehat{\mathcal{S}} := [\mathcal{S} \circ \mathbf{s}](\boldsymbol{\phi}) \tag{4.13}$$

where the dependence of $\mathbf{s}$ on $\boldsymbol{\phi}$ has been made explicit. We note that a phase-only formulation reduces the number of design variables by half while eliminating the N modulus constraints altogether.

4.1.3 Numerical Example

The definitions in the previous subsection can be combined to form six different nonlinear programming problems (programs). These programs, denoted $P_1 - P_6$, are listed in Table 4.1. The second column of the table describes the objective function, while the third and fourth columns describe the ACS and modulus constraints, respectively. Programs P_1 and P_4 represent a full waveform design (amplitude and phase), whereas P_3 and P_6 are true phase-only designs. Programs P_2 and P_5 are hybrids in that the real and imaginary parts are design separately, but they are constrained by the relationship $\mathrm{Re}\{s_n\}^2 + \mathrm{Im}\{s_n\}^2 = p_n$. Thus, P_2 and P_4 are essentially phase-only designs. We expect P_2 and P_4 to achieve the same gains as P_3 and P_6, respectively, while requiring more time to converge due to the additional constraints. We observe that P_1 has the most degrees of freedom, and is therefore

		Subject to		Number of	
ID	Maximize	ACS	Modulus	Design Variables	Constraints
P_1	$F(\mathbf{s})$	$\tilde{\mathcal{A}}(\mathbf{s}) \preceq \mathbf{0}$	$\mathcal{M}(\mathbf{s}) \preceq \mathbf{0}$	$2N$	$2N-1$
P_2	$F(\mathbf{s})$	$\tilde{\mathcal{A}}(\mathbf{s}) \preceq \mathbf{0}$	$\mathcal{M}(\mathbf{s}) = \mathbf{0}$	$2N$	$2N-1$
P_3	-$\widehat{F}(\phi)$	$\widehat{\mathcal{A}}(\phi) \preceq \mathbf{0}$	$-$	N	$N-1$
P_4	$F(\mathbf{s})$	$\mathcal{S}(\mathbf{s}) \leq 0$	$\mathcal{M}(\mathbf{s}) \preceq \mathbf{0}$	$2N$	$N+1$
P_5	$F(\mathbf{s})$	$\mathcal{S}(\mathbf{s}) \leq 0$	$\mathcal{M}(\mathbf{s}) = \mathbf{0}$	$2N$	$N+1$
P_6	-$\widehat{F}(\phi)$	$\widehat{\mathcal{S}}(\phi) \leq 0$	$-$	N	1

Table 4.1: Various Nonlinear Programs.

expected to provide the greatest gain. On the other hand, P_6 has half as many design variables and only one constraint. Thus, it is expected to converge to a solution fastest.

Each of the programs in Table 4.1 was solved using the MATLAB Optimization Toolbox [136] for the example interference process described in Sec. 3.3.2. In order to provide a fair comparison, each problem was solved using both the SQP algorithm [136, pp. 4-26] and the interior point algorithm [136, pp. 4-36]. However, only the results for the most efficient algorithm are reported. The algorithms were allowed to run for 10 minutes before termination, but they were also allowed to terminated early if a local optimum was found. Programs P_1 and P_2 were solved most efficiently by the IP algorithm using an initial barrier parameter of 1000. The remaining programs were solved more efficiently by SQP.

The ACS mask was constructed as described in Sec. 3.3.2 with the baseline waveform defined by

$$s_d(n) = \exp\left(j\left(\frac{\pi}{(N-1)}n^2 + \pi n\right)\right), \quad n = 0, \cdots, N-1 \tag{4.14}$$

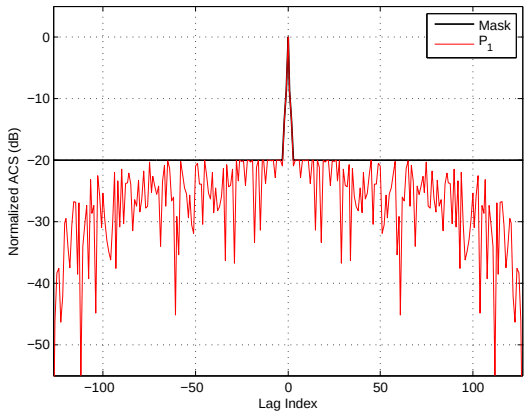
for $N = 128$, and a peak side lobe ratio (PSLR) of -20 dB. In order to provide a fair comparison between the direct and indirect ACS constraint methods, the similarity constant ϵ was chosen so that the final waveforms did not violate the constraint mask. This value

was found by trial-and-error to be $\epsilon = 0.01$. The ACSs of the optimized waveforms for each program are shown in Figs. 4.1 and 4.2. Notice that each design satisfies the ACS constraint mask.

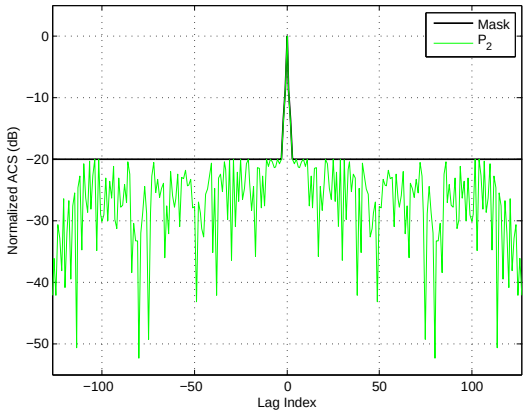
Figures 4.3 and 4.4 depict the evolution of gain over time for each program. Comparing program P_1 to P_4, P_2 to P_5, and P_3 to P_6, we observe that use of the similarity constraint significantly reduces convergence time, but at the cost of greatly reduced gain. Comparing P_2 to P_3 and P_5 to P_6, we notice a significant reduction in convergence time for phase-only design. Furthermore, comparing the gains of P_1 to P_3 and P_4 to P_6 suggests that amplitude modulation may play a negligible role in cost function reduction. As such, phase-only design may be a viable alternative to magnitude and phase design when modulus constraints are imposed.

4.1.4 Conclusion

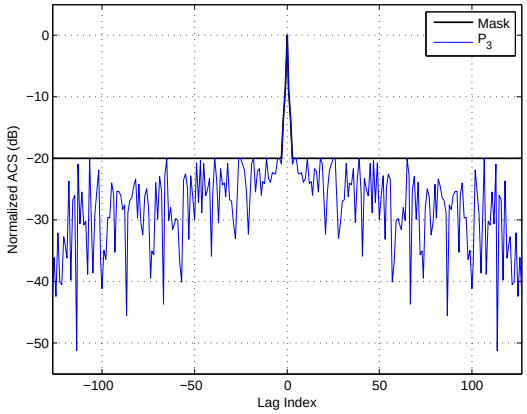
The results of this section suggest that the similarity constraint can be an ineffective means of constraining waveform ACS when modulus constraints are also imposed. The results also suggest that amplitude modulation may contribute only slightly to SINR maximization. As such, the phase-only design approach, which has fewer design variables and constraints, may be preferable. The IP and SQP algorithms used to obtain the results are extremely computationally complex. As such, more computationally efficient approaches should be investigated.



(a) ACS of P_1. Provided 39.86 dB of gain.

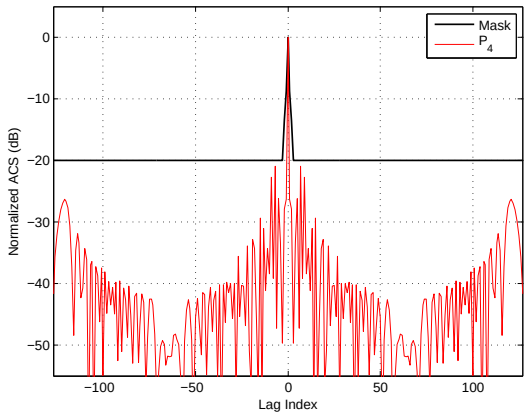


(b) ACS of P_2. Provided 38.26 dB of gain.

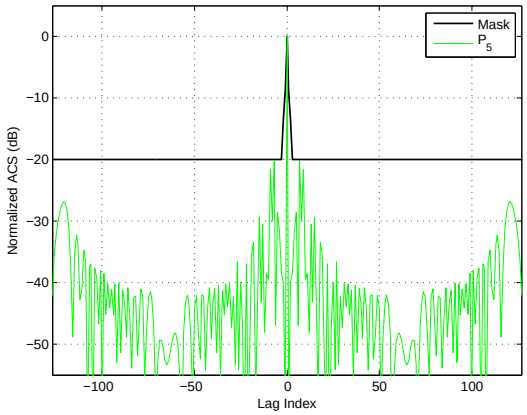


(c) ACS of P_3. Provided 38.24 dB of gain.

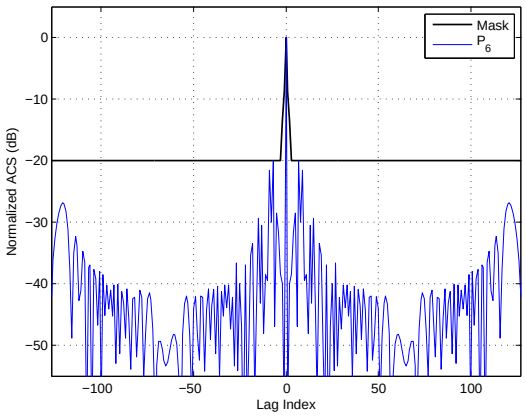
Figure 4.1: The -20 dB PSLR ACS mask is shown along with the squared magnitude of the final waveform ACS for programs (a) P_1 (b) P_2 and (c) P_3. P_1 provided 39.9 dB of gain above the LFM/MF solution. P_2 provided 38.3 dB, and P_3 provided 38.2 dB of gain.



(a) ACS of P_4 (ϵ=0.0100). Provided 2.49 dB of gain.



(b) ACS of P_5 (ϵ=0.0100). Provided 2.05 dB of gain.



(c) ACS of P_6 (ϵ=0.0100). Provided 2.05 dB of gain.

Figure 4.2: The -20 dB PSLR ACS mask is shown along with the squared magnitude of the final waveform ACS for programs (a) P_4 (b) P_5 and (c) P_6. P_4 provided 2.5 dB of gain above the LFM/MF solution. P_5 provided 2.0 dB, and P_6 provided 2.0 dB of gain.

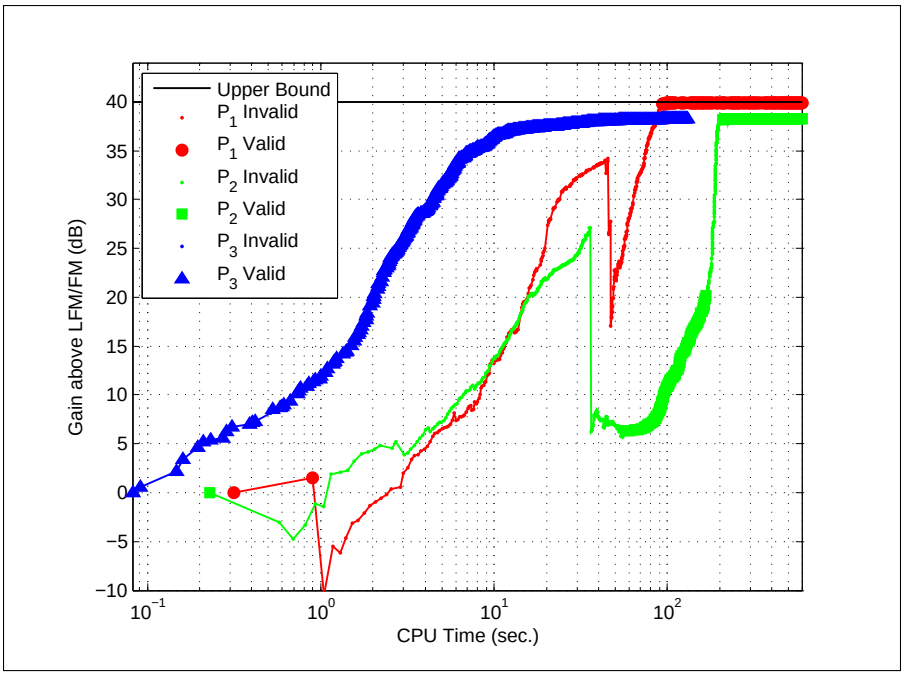


Figure 4.3: Learning curves for programs P_1-P_3.

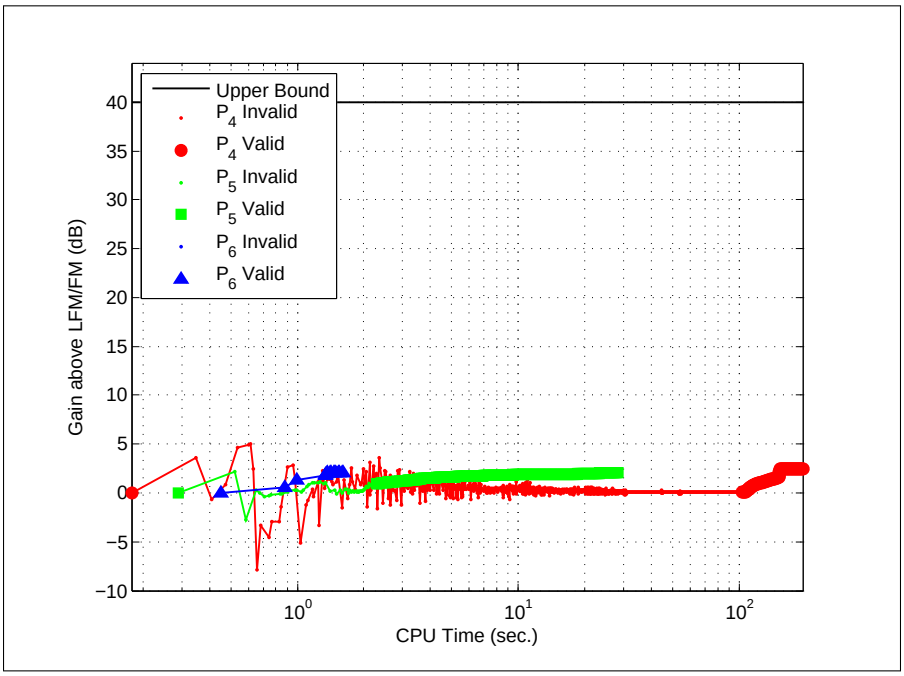


Figure 4.4: Learning curves for programs P_4-P_6.

4.2 NLFM for Waveform Optimization

In the previous section, we showed that a phase-only design can provide near optimal gain at a much lower computational cost, and indirectly constraining the waveform ACS via a similarity constraint can over restrict the feasible region of optimization. In this section, we consider another indirect approach that requires significantly fewer calculations than the direct ACS constraint approach while providing more gain than the similarity constraint method. Our approach is to represent the phase of the waveform by a parametrized approximation. We then relate this approximation to the ACS of the signal to demonstrate that the ACS can be indirectly constrained by limiting the magnitudes of the approximation parameters. The optimization is then performed with respect to these parameters.

4.2.1 Problem Formulation

Consider the scenario in which an attenuated and delayed version of the transmit signal ($\mathbf{s}$) is received in the presence of a zero-mean additive complex Gaussian random process, which is assumed to be independent of the transmitted signal. Note, this independence assumption is made for convenience, as our results can be applied without modification to the signal-dependent (i.e.., clutter/reverberation) case described in [28, 30, 79]. Assume that the received signal-plus-interference is processed by a correlation receiver, in which the receive filter is matched to the transmit signal [103, ch. 4]. Consider the functional $J : \mathbb{C}^N \rightarrow \mathbb{R}$ given by

$$J(\mathbf{s}) := \mathbf{s}^H \mathbf{K} \mathbf{s} \tag{4.15}$$

where $\mathbf{s} \in \mathbb{C}^N$ is the waveform to be optimized, and $\mathbf{K} \in \mathbb{C}^{N \times N}$ is the interference covariance matrix, which is assumed to be positive definite. It was shown in Ch. 3 that the deflection coefficient (i.e., post-filter signal-to-noise ratio) for the matched filter receiver is equal to $||\mathbf{s}||^4 / J(\mathbf{s})$. Thus, minimizing $J(\mathbf{s})$ maximizes the deflection coefficient, and

consequently, the probability of detection as well [103][106][132]. When subject only to the energy constraint $\|\mathbf{s}\|^2 = 1$, the minimizer of (4.15) is given by $\mathbf{s} = \boldsymbol{\phi}_0$, where $\boldsymbol{\phi}_0$ is the eigenvector of $\mathbf{K}$ corresponding to the minimum eigenvalue λ_0. Unfortunately, this prescription for the transmitted signal is not amenable to practical transceiver architectures, because $\boldsymbol{\phi}_0$ will not possess, in general, the required magnitude function (i.e., constant modulus). Moreover while $\boldsymbol{\phi}_0$ minimizes J, its autocorrelation sequence may suffer from high side lobes and a broadened or fractured main lobe. We propose to design a phase-only waveform, thereby automatically satisfying any magnitude function requirement, and we propose to indirectly control the ACS by constraining the waveform parameters that minimize J. It should be noted that this approach can be easily extended to design waveforms having any magnitude function, not just constant modulus.

We select the linear frequency-modulated (LFM) pulse defined by

$$s_L(n) = \exp\left(j(\gamma n^2 + \sigma n)\right), \quad n = 0, \cdots, N-1 \tag{4.16}$$

as our starting point because of the spectral flatness it provides. Nonlinear perturbation of the phase of this waveform will provide spectral selectivity by sweeping through some frequencies faster or slower than others. This will also affect the ACS shape. We note that the idea of nonlinear frequency modulation (NLFM) is by no means new, having already been employed for applications such as side lobe control [109].

Following the paradigm of [109, Section 10.4], a Fourier series phase perturbation is selected, giving the signal model

$$s(n) = s_L(n)\exp\left(j\sum_{k=1}^{M-1} a_k \sin(\omega_k n) + b_k \cos(\omega_k n)\right) \tag{4.17}$$

where $\omega_k = \pi k/M$. We can now derive the ACS of (4.17) by introducing limited approxi-

mations. By definition, the ACS of (4.17), evaluated at lag l, is given by

$$r(l) := \sum_{n=0}^{N-l-1} s(n)s^*(n-l) \tag{4.18}$$

Expanding this expression, applying standard trigonometric identities, and simplifying results in

$$\begin{aligned} r(l) = & \sum_{n=0}^{N-l-1} \exp\left(j\left(2\gamma nl + \phi\right)\right) \\ & \times \exp\left(j2\sum_{k=0}^{M-1} \sin\left(\frac{\omega_k l}{2}\right)\left[a_k \cos\left(\omega_k\left(n+\frac{l}{2}\right)\right) - b_k \sin\left(\omega_k\left(n+\frac{l}{2}\right)\right)\right]\right) \end{aligned}$$

where $\phi = \gamma l^2 + \sigma l$. When the argument of the second exponential is small, we can approximate the above by

$$\begin{aligned} \widehat{r}(l) = & \sum_{n=0}^{N-l-1} \exp\left(j\left(2\gamma nl + \phi\right)\right) \\ & \times \left(1 + 2j\sum_{k=0}^{M-1} \sin\left(\frac{\omega_k l}{2}\right)\left[a_k \cos\left(\omega_k\left(n+\frac{l}{2}\right)\right) - b_k \sin\left(\omega_k\left(n+\frac{l}{2}\right)\right)\right]\right) \end{aligned}$$

Rearranging the summations, applying Euler's identity and the geometric series formula, and simplifying yields

$$\widehat{r}(l) = \zeta(l) D_l\left(\gamma l\right) + \zeta(l) \sum_{k=0}^{M-1} \left\{\alpha_k D_l\left(\gamma l + \frac{\omega_k}{2}\right) + \beta_k D_l\left(\gamma l - \frac{\omega_k}{2}\right)\right\} \tag{4.19}$$

where

$$\zeta(l) = \exp\left(j(\gamma l(N-1) + \sigma l)\right) \tag{4.20}$$

$$D_l\left(\Omega\right) = \frac{\sin\left((N-l)\Omega\right)}{\sin\left(\Omega\right)} \tag{4.21}$$

and

$$\alpha_k = \sin\left(\frac{\omega_k l}{2}\right) \exp\left(j\left(\gamma l + \frac{\omega_k l}{2}\right)(N-1)\right)(a_k j - b_k) \tag{4.22}$$

$$\beta_k = \sin\left(\frac{\omega_k l}{2}\right) \exp\left(j\left(\gamma l - \frac{\omega_k l}{2}\right)(N-1)\right)(a_k j + b_k) \tag{4.23}$$

The first term in (4.19) can be recognized as the ACS of the LFM (s_L). Thus, the approximate increase in side lobes is given by

$$\tilde{r}(l) = \zeta(l) \sum_{k=0}^{M-1} \left\{ \alpha_k D_l\left(\gamma l + \frac{\omega_k}{2}\right) + \beta_k D_l\left(\gamma l - \frac{\omega_k}{2}\right) \right\}, \tag{4.24}$$

and we see that by constraining the Fourier coefficients a_k and b_k to be small, we constrain the ACS error to be small as well.

While the result of (4.24) is informative, it does not provide an exact method of predicting increased side lobes in the presence of sinusoidal phase perturbations. We therefore define the waveform optimization problem:

$$\min_{\{a_k, b_k\}} \ \mathbf{s}^H \mathbf{K} \mathbf{s} \tag{4.25a}$$

$$\text{subject to} \quad |a_k| < a_{max}, \quad k = 0, \ldots, N-1 \tag{4.25b}$$

$$|b_k| < a_{max}, \quad k = 0, \ldots, N-1 \tag{4.25c}$$

where the waveform $\mathbf{s}$ is defined by (4.17), and a_{max} controls the maximum allowable side lobe increase. For a given number of coefficients ($2M$), the parameter a_{max} controls the size of the feasible region of optimization. As a_{max} is increased, greater nonlinear phase perturbations will be permitted. This will allow increased performance improvement, but this improvement will be at the expense of higher ACS side lobes. The optimization problem defined in (4.25) can be initialized with $a_k, b_k = 0$ for all k, and the optimization can be performed efficiently by any number of constrained optimization algorithms such

as sequential quadratic programming (SQP) or interior point methods [131]. An efficient method for calculating the gradient of the cost function with respect to the parameter vector $[a_0 \ \ldots \ a_N \ b_1 \ \ldots \ b_N]$ is provided in the Appendix.

4.2.2 Simulation Results

In order to illustrate the effectiveness of the NLFM-optimized approach (4.25), we compare its performance to that of the direct and similarity constrained approaches for the example interference process shown in Fig. 4.5. This interference process can be generated by filtering circularly symmetric complex-valued white Gaussian noise with the filter

$$H(z) = \frac{1}{(1 - 1.5z^{-1} + 0.7z^{-2})^4} \tag{4.26}$$

(i.e., an autoregressive process), and then adding the result to another circularly symmetric complex-valued white Gaussian random process such that the resulting interference-to-noise ratio is 40 dB, and the resulting signal-to-interference-plus-noise ratio is -5 dB. However, for the following examples, we assume perfect knowledge of the interference covariance matrix $\mathbf{K}$, which can be calculated in closed form [137].

If we define the transmit signal as $s_n := \exp(j\phi_n)$, then the directly constrained design can be found by solving the optimization problem

$$\min_{\{\phi_n\}} \ \mathbf{s}^H \mathbf{K} \mathbf{s} \tag{4.27a}$$

$$\text{subject to } \ |r(l)|^2 \leq m_l, \ \ l = 0, \cdots, N-1 \tag{4.27b}$$

where m_l is the allowable squared magnitude of the waveform ACS at lag l. The actual values for the mask will be discussed later. The similarity constrained design can be found

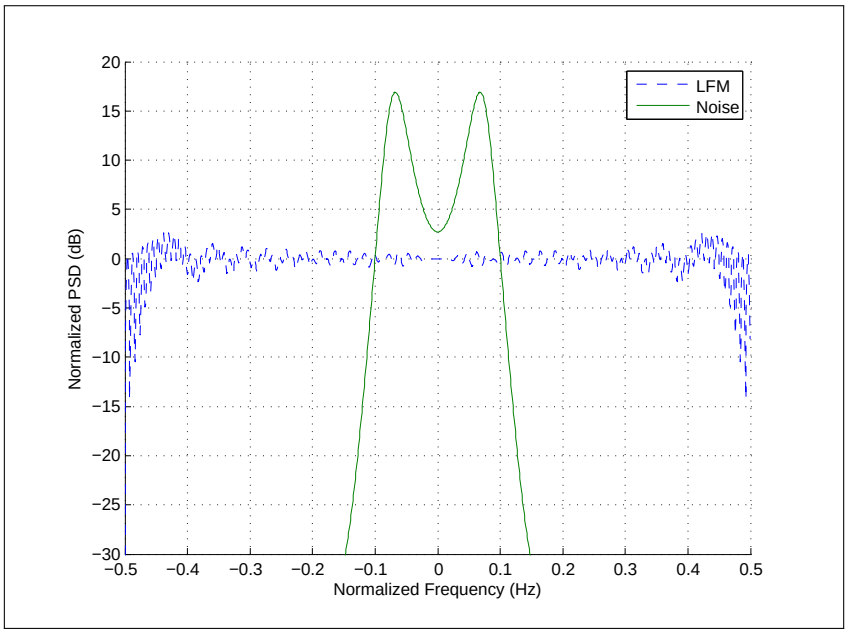


Figure 4.5: Power spectral density of interference and baseline LFM.

by solving

$$\min_{\{\phi_n\}} \ \mathbf{s}^H \mathbf{K} \mathbf{s} \tag{4.28a}$$

$$\text{subject to } \ ||\mathbf{s}_L - \mathbf{s}||^2 \leq \epsilon \tag{4.28b}$$

where $\mathbf{s}_L$ is a waveform with a desirable ACS, and ϵ is the similarity constant. Designs (4.27) and (4.28) were discussed in the previous section. For each design, the length of the waveform was specified to be $N = 128$. The template waveform $\mathbf{s}_L$ in (4.28) was chosen to be the LFM defined in (4.16), with $\gamma = \pi/(N-1)$ and $\sigma = \pi$. Both indirect approaches require user parameters: M and a_{max} for the NLFM approach, and ϵ for the similarity approach. In order to make a fair comparison between the three designs, we set $M = N/2$, and then manually found the largest a_{max} and ϵ such that the ACSs of the optimized waveforms did not violate the constraint mask m. This way, each design was given the same number of design variables and the same ACS constraints.

Algorithm performance is measured in two ways. The first is the gain of the solution

with respect to the initial waveform s_L. This value is given by

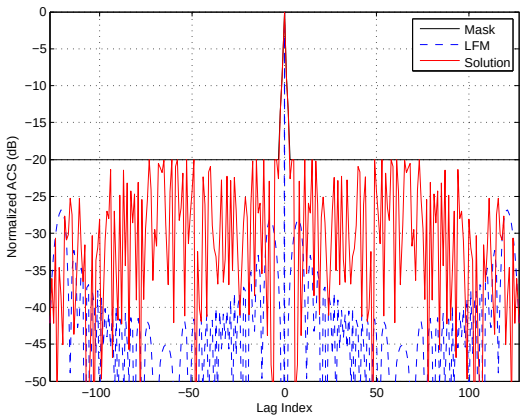
$$G(\mathbf{s}) := \frac{J(\mathbf{s}_L)}{J(\mathbf{s})} = \frac{\mathbf{s}_L^H \mathbf{K} \mathbf{s}_L}{\mathbf{s}^H \mathbf{K} \mathbf{s}} \tag{4.29}$$

For the spectrum shown in Fig. 4.5, the maximum achievable gain for a unit energy waveform, without regard for ACS or modulus constraints, is

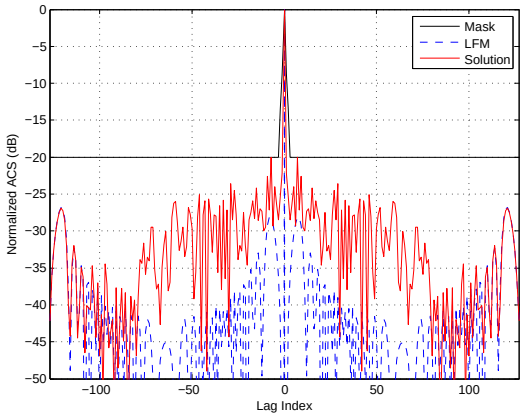
$$G(\boldsymbol{\phi}_0) = \frac{\mathbf{s}_L^H \mathbf{K} \mathbf{s}_L}{\lambda_0} = 40 \text{ dB}. \tag{4.30}$$

The second measure of algorithm performance is the time required to compute the solution. Each optimization problem was solved using the MATLAB Optimization Toolbox [136]. In order to make a fair comparison, each problem was solved using the sequential quadratic programming (SQP) and the interior-point algorithms, and an effort was made in each case to find the algorithm options that resulted in the fastest convergence time. Problem (4.27) was solved the fastest using the SQP algorithm [136, pp. 4-26], whereas (4.28) and (4.25) were solved more efficiently by the interior-point algorithm [136, pp. 4-36]. Only the results from the fastest algorithm for each problem are reported.

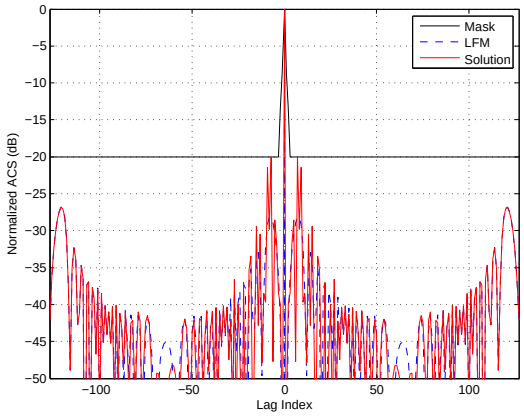
The first comparison between the three designs was performed for the ACS mask depicted in Figs. 4.6(a)-4.6(c), which also show the final solution ACSs. This mask limits the ACS peak side lobe ratio (PSLR) to -20 dB, and we note that each design satisfies the ACS constraint. The evolution of gain over time (seconds) is shown in Fig. 4.7 for each optimization problem. Notice that the NLFM approach provides higher gain than the similarity constraint at each epoch. However, the direct approach begins to provide higher gain at approximately 1.5 seconds, continuing to improve until reaching a substantially higher gain. This example shows that the NLFM-optimized approach can be beneficial for scenarios in which little time is available to compute the solution. This may occur in applications involving nonstationary interference that can be assumed stationary over a short interval of time.



(a) Directly constrained solution. Run time of 30.33 seconds resulted in 37.73 dB gain and -20.00 dB PSLR.



(b) NLFM solution (M=64, a_{max}=0.0588). Run time of 2.88 seconds resulted in 5.31 dB gain and -20.01 dB PSLR.



(c) Similarity constrained solution (ϵ=0.0103). Run time of 1.21 seconds resulted in 2.08 dB gain and -20.03 dB PSLR.

Figure 4.6: The -20 dB PSLR ACS mask is shown along with the squared magnitude of the final waveform ACS for all three methods.

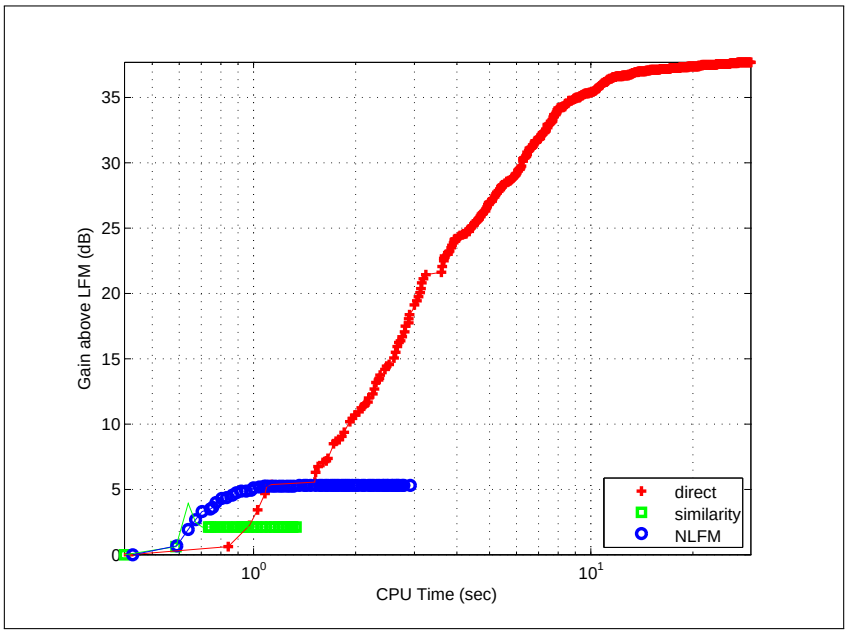


Figure 4.7: Gain versus CPU time for the -20 PSLR example.

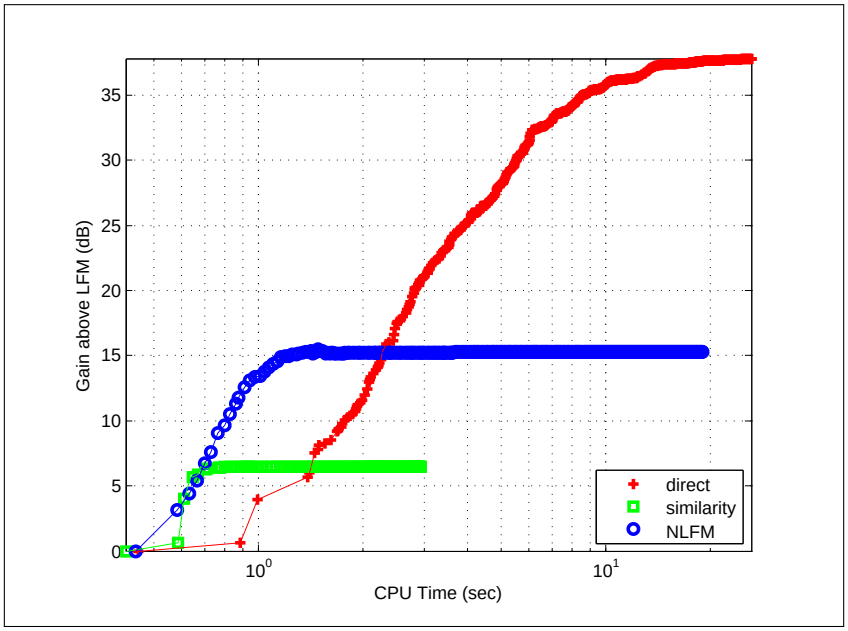
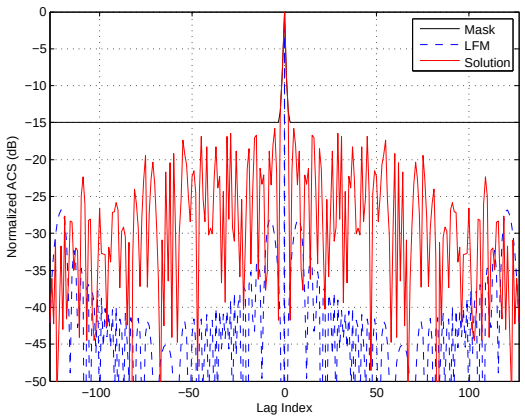


Figure 4.8: Gain versus CPU time for the -15 PSLR example.

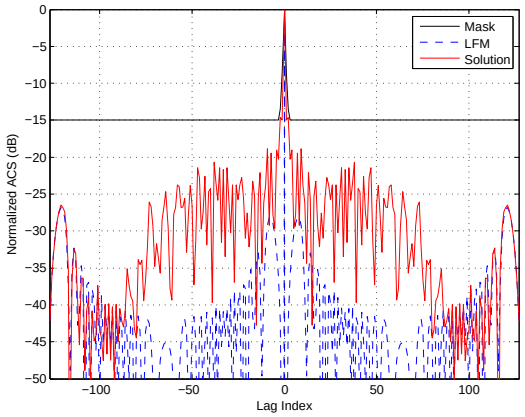
In order to illustrate the trade between gain and PSLR that occurs with increasing a_{max}, we consider the case in which the PSLR was constrained to be less than -15 dB. The ACS mask for this example is shown in Figs. 4.9(a)-4.9(c), which also depict the final solution ACSs. Again, we note that each solution satisfies the ACS mask constraint. The evolution of gains over time for this example is shown in Fig. 4.8. The increase in gain resulting from the relaxed ACS constraints was negligible for the direct approach, but significant for both indirect approaches. As before, the NLFM approach provides higher gain than the similarity approach at each epoch. However, the direct ACS begins to provide higher gain at approximately 2.5 seconds, continuing until a substantially higher gain is achieved. This example illustrates that the NLFM approach provides a mechanism for trading between gain and PSLR.

4.2.3 Conclusions

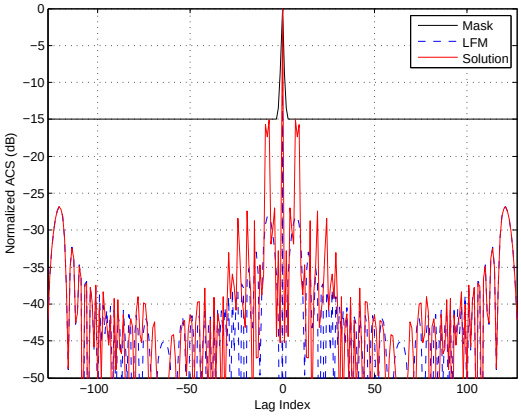
Practical waveform design algorithms must seek to improve system performance while maintaining good ACS properties and satisfying modulus requirements for implementation. In this section, we described a phase-only waveform optimization process that allows simple indirect constraints to be placed on ACS side lobe levels, providing a straightforward means to trade gain for increased side lobes. This approach outperforms the similarity constraint method in terms of gain and convergence time. Furthermore, it can provide greater gain than the directly constrained approach in applications for which only a minimal amount of time is available to compute the solutions, but as the allowed computation time is extended, the NLFM performance gain plateaus due to the over constrained feasibility region.



(a) Directly constrained solution. Run time of 26.19 seconds resulted in 37.81 dB gain and -15.76 dB PSLR.



(b) NLFM solution (M=64, a_{max}=0.1250). Run time of 18.96 seconds resulted in 15.31 dB gain and -14.65 dB PSLR.



(c) Similarity constrained solution (ϵ=0.0710). Run time of 2.94 seconds resulted in 6.46 dB gain and -15.06 dB PSLR.

Figure 4.9: The -15 dB PSLR ACS mask is shown along with the squared magnitude of the final waveform ACS for all three methods. 105

Chapter 5

Closing Remarks

In this chapter, we summarize the contributions made by this dissertation (5.1) and elaborate on how this research can be extended (5.2).

5.1 Contributions

Prior to our work, the waveform optimization literature consisted of incomplete problem formulations that did not capture all of the constraints that would actually be imposed on the transmit waveform by a radar system. Thus, for most practical scenarios, those results *cannot be used to design an actual ATx system*. We have addressed this shortcoming, by 1) identifying the critical constraints (i.e., waveform modulus and ambiguity function), and 2) providing at least baseline solutions to the resulting optimization problems. Thus, our results represent significant advances in ATx technology – advances that move this technology toward implementability. We will now discuss the individual contributions of this dissertation in more detail.

Modulus Constraints: (Ch. 2.) It has long been appreciated that the waveform optimization algorithms found in the literature do not incorporate constant modulus constraints. These constraints are non-convex, and are therefore difficult to satisfy in otherwise convex formulations. As a result, many works assume that a linear power amplifier is used instead

of a nonlinear amplifier; this is the case for the state of the art solutions of [33] and [28]. However, we demonstrated that the waveform modulus must still be constrained when a linear amplifier is used. Further, we have shown that any algorithm that does not do so may actually *degrade system performance*. This result may impact the field of waveform optimization by leading future algorithm designs to account for modulus constraints.

Modulus Constraint Satisfaction: (Ch. 2.) Having established the inextricable nature of modulus constraints, we went on to consider how the state of the art in waveform optimization could be extended to satisfy them. Our contribution was to 1) establish the connection between this problem and the problem of phase retrieval, 2) make advances in the convergence analysis of applicable algorithms, 3) derive a novel algorithm for a specific case (when the number of degrees of freedom in the time domain exceeds those in the frequency domain), and 4) demonstrate that the alternating projections approach is a computationally efficient alternative to "parameter free" nonlinear programming methods.

Multi-Target Detection: (Ch. 3.) Heretofore, the waveform optimization literature only considered the detection of a *single* target. When multiple targets are present at unknown ranges/velocities, which is by far the most common radar scenario, the algorithms in the literature are not applicable because they do not consider the ambiguity function of the transmit waveform. Our contribution was to 1) formulate the waveform optimization problem for a variety of receiver architectures, 2) introduce the direct method of constraining the waveform ambiguity/cross-ambiguity function, and 3) provide an investigation of applicable nonlinear programming algorithms. Our solutions are the state of the art for the design of interference-suppressing waveforms in the multi-target detection problem.

Efficient Formulations, Similarity Constraint: (Ch. 4.) The direct ambiguity function constraint approach is computationally demanding. Thus, additional contributions of this dissertation are that we 1) introduced alternate formulations involving phase-only designs and similarity constraints on phase-only designs, and 2) we demonstrated that the similarity constraint is an inefficient means for constraining the waveform ambiguity func-

tion. This result is significant because the similarity constraint (for magnitude and phase design) has been proposed in other works without considering the effect on optimality.

Efficient Formulations, NLFM: (Ch. 4.) Another contribution of this dissertation was that we introduced a novel approach to constraining a waveform's ambiguity function via a parametrization of nonlinear frequency modulation. This approach provides gains greater than the similarity constraint, but at a computational cost much less than than the direct constraint approach.

5.2 Future Research

There are perhaps two issues that must be addressed before ATx technology can be implemented in a real system. The first involves constraining the bandwidth of the transmit signal. Our nonlinear programming approach can be extended to handle these constraints in a straightforward manner. Secondly, many applications will require the fast computation of solutions. It is not yet clear if our efficient formulations are fast enough. This will be known only after a specific application has been chosen, and a prototype system is designed.

We note that clutter has not been explicitly considered in this work. It is a straightforward matter to augment our cost functions to account for the LTI clutter model that is prevalent in the literature. However, if the delay-frequency distribution of the clutter is known, then the direct ambiguity function constraint method can be used to make the AF response low in the portion of the delay-frequency plane occupied by the clutter. This approach should be investigated in future work.

Appendix A

Appendix for Chapter 2

A.1 Relevant Proofs

Theorem A.1-1 (Point-set distance is nonincreasing for alternating projections)**.** *Let S and T be subspaces of metric space (X, d) with projections Π_S and Π_T, respectively. Assume mapping $P_S : X \to S$ and $P_T : X \to T$ are well defined such that $P_S(\mathbf{x}) \in \Pi_S(\mathbf{x})$ and $P_T(\mathbf{x}) \in \Pi_S(\mathbf{x})$ for all $\mathbf{x} \in X$, and define the sequence $\{\mathbf{s}_k\}_{k=0}^{\infty}$ such that $\mathbf{s}_{k+1} = (P_S \circ P_T)(\mathbf{s}_k)$ for $k > 0$. Then, $d(\mathbf{s}_{k+1}, T) \leq d(\mathbf{s}_k, T)$.*

Proof. Define the sequence $\{\mathbf{t}_k\}_{k=0}^{\infty}$ by $\mathbf{t}_k = P_T(\mathbf{s}_k)$, and note that $\mathbf{s}_{k+1} = P_S(\mathbf{t}_k)$. Then, by the properties of a projection, and the symmetry of the metric d, we have

$$
\begin{aligned}
d(\mathbf{s}_{k+1}, T) &\leq d(\mathbf{s}_{k+1}, \mathbf{t}_k) && \text{(A.1)} \\
&= d(\mathbf{t}_k, \mathbf{s}_{k+1}) && \text{(A.2)} \\
&= d(\mathbf{t}_k, S) && \text{(A.3)} \\
&\leq d(\mathbf{t}_k, \mathbf{s}_k) && \text{(A.4)} \\
&= d(\mathbf{s}_k, \mathbf{t}_k) && \text{(A.5)} \\
&= d(\mathbf{s}_k, T) && \text{(A.6)}
\end{aligned}
$$

□

Theorem A.1-2. *Let set A be defined by (2.13) with projection operator Π_A. Let $P_A : \mathbb{C}^N \to A$ be defined by (2.15). Then $P_A(\mathbf{x}) \in \Pi_A(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{C}^N$.*

Proof. Note that $||\mathbf{a}||^2 = ||\mathbf{p}||^2$ for all $\mathbf{a} \in A$. Thus, for any arbitrary but fixed $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{a} \in A$, we can write

$$||\mathbf{x} - \mathbf{a}||^2 = ||\mathbf{x}||^2 + ||\mathbf{p}||^2 - 2\mathrm{Re}\left\{\mathbf{x}^H\mathbf{a}\right\} \tag{A.7}$$

Therefore,

$$\Pi_A(\mathbf{x}) := \arg\inf_{\mathbf{a}\in A} ||\mathbf{x} - \mathbf{a}||^2 = \arg\sup_{\mathbf{a}\in A} \mathrm{Re}\left\{\mathbf{x}^H\mathbf{a}\right\} \tag{A.8}$$

By definition of A, we have

$$\mathrm{Re}\left\{\mathbf{x}^H\mathbf{a}\right\} = \sum_{n=0}^{N-1} p_n |x_n| \cos\left(\angle x_n - \angle a_n\right) \leq \sum_{n=0}^{N-1} p_n |x_n| \tag{A.9}$$

Thus, the minimum distance is achieved $P_A(\mathbf{x})$ defined by $[P_A(\mathbf{x})]_n = p_n \exp\left(j\angle \mathbf{x}_n\right)$. (Note that $\Pi_A(\mathbf{x})$ is set valued in general because the phase of $\mathbf{a} \in \Pi_A(\mathbf{x})$ can take on any value wherever $x_n = 0$. Thus, $P_A \neq \Pi_A$.) □

Theorem A.1-3. *Let set $\widetilde{B}$ be defined by (2.14) with projection operator $\Pi_{\widetilde{B}}$. Let $P_{\widetilde{B}} : \mathbb{C}^M \to \widetilde{B}$ be defined by (2.16). Then $P_{\widetilde{B}}(\mathbf{x}) \in \Pi_{\widetilde{B}}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{C}^M$.*

Proof. Replace A with $\widetilde{B}$, N with M and $\mathbf{p}$ with $\mathbf{q}$ in the proof of Thm. A.1-2. □

Theorem A.1-4. *Let $A \subseteq \mathbb{C}^N$ have projection Π_A, and assume there exists a mapping $P_A : \mathbb{C}^N \to A$ such that $P_A(\mathbf{x}) \in \Pi_A(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{C}^N$. Let $\mathbf{U} \in \mathbb{C}^{M\times N}$ be such that $\mathbf{U}^H\mathbf{U} = \mathbf{I}$, and denote the image of A under $\mathbf{U}$ by $\widehat{A}$. Then the mapping $P_{\widehat{A}} := \mathbf{U}P_A\mathbf{U}^H$ is such that $P_{\widehat{A}}(\tilde{\mathbf{x}}) \in \Pi_{\widehat{A}}(\tilde{\mathbf{x}})$ for all $\tilde{\mathbf{x}} \in \mathbb{C}^M$.*

Proof. Every $\tilde{\mathbf{x}} \in \mathbb{C}^M$ has a unique orthogonal decomposition $\tilde{\mathbf{x}} = \widehat{\mathbf{x}} + \tilde{\mathbf{x}}_n$ where $\widehat{\mathbf{x}} = \mathbf{U}\mathbf{U}^H\tilde{\mathbf{x}} \in \mathcal{R}(\mathbf{U})$ and $\tilde{\mathbf{x}}_n \in \mathcal{N}\left(\mathbf{U}^H\right)$. Let $\tilde{\mathbf{x}} \in \mathbb{C}^M$ and $\widehat{\mathbf{a}} \in \widehat{A}$ be arbitrary but fixed. Then,

$$||\tilde{\mathbf{x}} - \widehat{\mathbf{a}}||^2 = ||\tilde{\mathbf{x}}_n||^2 + ||\widehat{\mathbf{x}} - \widehat{\mathbf{a}}||^2 \tag{A.10}$$

$$= ||\tilde{\mathbf{x}}_n||^2 + \left|\left|\mathbf{U}(\mathbf{U}^H\tilde{\mathbf{x}} - \mathbf{U}^H\widehat{\mathbf{a}})\right|\right|^2 \tag{A.11}$$

$$= ||\tilde{\mathbf{x}}_n||^2 + \left|\left|\mathbf{U}^H\mathbf{x} - \mathbf{U}^H\widehat{\mathbf{a}}\right|\right|^2 \tag{A.12}$$

The point $\mathbf{U}^H\mathbf{x}$ must be at least as close to its projection onto A than it is to the point $\mathbf{U}^H\widehat{\mathbf{a}} \in A$. Thus,

$$||\tilde{\mathbf{x}} - \widehat{\mathbf{a}}||^2 \geq ||\tilde{\mathbf{x}}_n||^2 + \left|\left|\mathbf{U}^H\tilde{\mathbf{x}} - P_A(\mathbf{U}^H\tilde{\mathbf{x}})\right|\right|^2 \tag{A.13}$$

$$= ||\tilde{\mathbf{x}}_n||^2 + \left|\left|\mathbf{U}(\mathbf{U}^H\tilde{\mathbf{x}} - P_A(\mathbf{U}^H\tilde{\mathbf{x}}))\right|\right|^2 \tag{A.14}$$

$$= ||\tilde{\mathbf{x}}_n||^2 + \left|\left|\widehat{\mathbf{x}} - \mathbf{U}P_A(\mathbf{U}^H\tilde{\mathbf{x}})\right|\right|^2 \tag{A.15}$$

$$= \left|\left|\tilde{\mathbf{x}}_n + \widehat{\mathbf{x}} - \mathbf{U}P_A(\mathbf{U}^H\tilde{\mathbf{x}})\right|\right|^2 \tag{A.16}$$

$$= \left|\left|\tilde{\mathbf{x}} - P_{\widehat{A}}(\tilde{\mathbf{x}})\right|\right|^2 \tag{A.17}$$

Thus, $\left|\left|\widehat{\mathbf{x}} - P_{\widehat{A}}(\tilde{\mathbf{x}}))\right|\right| \leq ||\tilde{\mathbf{x}} - \widehat{\mathbf{a}}||$, and since $P_{\widehat{A}}(\tilde{\mathbf{x}}) \in \widehat{A}$, we have $P_{\widehat{A}}(\tilde{\mathbf{x}}) \in \Pi_{\widehat{A}}(\tilde{\mathbf{x}})$. □

Theorem A.1-5. *Let $\widetilde{B} \subseteq \mathbb{C}^M$ have projection $\Pi_{\widetilde{B}}$, and assume there exists a mapping $P_{\widetilde{B}} : \mathbb{C}^M \to \widetilde{B}$ such that $P_{\widetilde{B}}(\tilde{\mathbf{x}}) \in \Pi_{\widetilde{B}}(\tilde{\mathbf{x}})$ for all $\tilde{\mathbf{x}} \in \mathbb{C}^M$. Let $\mathbf{V} \in \mathbb{C}^{M\times N}$ be such that $\mathbf{V}\mathbf{V}^H = \mathbf{I}$. Let $B \subseteq \mathbb{C}^N$ denote the pre-image of $\widetilde{B}$, that is, those $\mathbf{x} \in \mathbb{C}^N$ with $\mathbf{V}\mathbf{x} \in \widetilde{B}$. Then the mapping $P_B : \mathbb{C}^N \to B$ defined by $P_B(\mathbf{x}) := \mathbf{V}^H P_{\widetilde{B}}(\mathbf{V}\mathbf{x}) + (\mathbf{I} - \mathbf{V}^H\mathbf{V})\mathbf{x}$ is such that $P_B(\mathbf{x}) \in \Pi_B(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{C}^N$.*

Proof. (Note, this proof is similar to A.1-4, but not equivalent.) Every $\mathbf{x} \in \mathbb{C}^N$ has a unique orthogonal decomposition $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$ where $\mathbf{x}_r = \mathbf{V}^H\mathbf{V}\mathbf{x} \in \mathcal{R}\left(\mathbf{V}^H\right)$ and $\mathbf{x}_n \in \mathcal{N}(\mathbf{V})$.

Let $\mathbf{x} \in \mathbb{C}^N$ and $\mathbf{b} \in B$ be arbitrary but fixed. Then,

$$||\mathbf{b} - \mathbf{x}||^2 = ||\mathbf{b}_r - \mathbf{x}_r||^2 + ||\mathbf{b}_n - \mathbf{x}_n||^2 \tag{A.18}$$

$$\geq ||\mathbf{b}_r - \mathbf{x}_r||^2 \tag{A.19}$$

$$= \left|\left|\mathbf{V}^H(\mathbf{V}\mathbf{b} - \mathbf{V}\mathbf{x})\right|\right|^2 \tag{A.20}$$

$$= ||\mathbf{V}\mathbf{b} - \mathbf{V}\mathbf{x}||^2 \tag{A.21}$$

By definition, $\mathbf{V}\mathbf{b} \in \widetilde{B}$. Thus,

$$||\mathbf{b} - \mathbf{x}||^2 \geq \left|\left|P_{\widetilde{B}}(\mathbf{V}\mathbf{x}) - \mathbf{V}\mathbf{x}\right|\right|^2 \tag{A.22}$$

$$= \left|\left|\mathbf{V}^H(P_{\widetilde{B}}(\mathbf{V}\mathbf{x}) - \mathbf{V}\mathbf{x})\right|\right|^2 \tag{A.23}$$

$$= \left|\left|\mathbf{V}^H P_{\widetilde{B}}(\mathbf{V}\mathbf{x}) - \mathbf{V}^H\mathbf{V}\mathbf{x}\right|\right|^2 \tag{A.24}$$

$$= \left|\left|\mathbf{V}^H P_{\widetilde{B}}(\mathbf{V}\mathbf{x}) + (\mathbf{I} - \mathbf{V}^H\mathbf{V})\mathbf{x} - \mathbf{x}\right|\right|^2 \tag{A.25}$$

$$= ||P_B(\mathbf{x}) - \mathbf{x}||^2 \tag{A.26}$$

Thus, $||\mathbf{x} - P_B(\mathbf{x}))|| \leq ||\mathbf{x} - \mathbf{b}||$, and since $P_B(\mathbf{x}) \in B$, we have $P_B(\mathbf{x}) \in \Pi_B(\mathbf{x})$. □

Theorem A.1-6. *Let sets $\widehat{A}$ and $\widetilde{B}$ be defined be defined by (2.17) and (2.14), respectively, and define the mappings $P_{\widehat{A}}$ and $P_{\widetilde{B}}$ by (2.18) and (2.16), respectively. Let J be defined by (2.23), and assume $\mathbf{U}^H\mathbf{U} = \mathbf{I}_N$. Then, $d^2(\widehat{\mathbf{a}}, \widetilde{B}) = J(\mathbf{U}^H\widehat{\mathbf{a}})$ for all $\widehat{\mathbf{a}} \in \widehat{A}$.*

Proof. Let $\widehat{\mathbf{a}} \in \widehat{A}$ be arbitrary but fixed. Then, $\widehat{\mathbf{a}} = \mathbf{U}\mathbf{U}^H\mathbf{a}$, and we have

$$d^2(\widehat{\mathbf{a}}, \widetilde{B}) := \left|\left|\widehat{\mathbf{a}} - P_{\widetilde{B}}(\widehat{\mathbf{a}})\right|\right|^2 \tag{A.27}$$

$$= \sum_{m=0}^{M-1} \left|(|\widehat{a}_m| - q_m)\exp\left(j\angle a_m\right)\right|^2 \tag{A.28}$$

$$= \sum_{m=0}^{M-1} \left||\widehat{a}_m| - q_m\right|^2 \tag{A.29}$$

$$= \left|\left||\widehat{\mathbf{a}}| - \mathbf{q}\right|\right|^2 \tag{A.30}$$

$$= J(\mathbf{U}^H\widehat{\mathbf{a}}) \tag{A.31}$$

□

Theorem A.1-7. *Let sets A and B be defined be defined by (2.13) and (2.21), respectively, and define the mappings P_A and P_B by (2.15) and (2.22), respectively. Let J be defined by (2.23), and assume $\mathbf{V}\mathbf{V}^H = \mathbf{I}_M$. Then, $d^2(\mathbf{a}, B) = J(\mathbf{a})$ for all $\mathbf{a} \in A$.*

Proof. Let $\mathbf{a} \in A$ be arbitrary but fixed. Then,

$$d^2(\mathbf{a}, B) := \left|\left|\mathbf{V}^H P_{\widetilde{B}}(\mathbf{V}\mathbf{a}) + (\mathbf{I} - \mathbf{V}^H\mathbf{V})\mathbf{a} - \mathbf{a}\right|\right|^2 \tag{A.32}$$

$$= \left|\left|\mathbf{V}^H P_{\widetilde{B}}(\mathbf{V}\mathbf{a}) - \mathbf{V}^H\mathbf{V}\mathbf{a}\right|\right|^2 \tag{A.33}$$

$$= \left|\left|\mathbf{V}^H (P_{\widetilde{B}}(\mathbf{V}\mathbf{a}) - \mathbf{V}\mathbf{a})\right|\right|^2 \tag{A.34}$$

$$= \left|\left|P_{\widetilde{B}}(\mathbf{V}\mathbf{a}) - \mathbf{V}\mathbf{a}\right|\right|^2 \tag{A.35}$$

$$= \left|\left||\widehat{\mathbf{a}}| - \mathbf{q}\right|\right|^2 \tag{A.36}$$

$$= J(\mathbf{a}) \tag{A.37}$$

where (A.36) follows from (A.27)-(A.30) in Thm. A.1-6. □

A.2 Valid DTFT Operators

A.2.1 A General Discrete-Time Fourier Transform

In the spectrum shaping problem for radar waveform optimization, a desired Fourier transform magnitude (FTM) is specified at a finite number of frequencies. As such, the transformation matrix $\mathbf{F}$ should compute the discrete-time Fourier transform of the input vector at just those frequencies of interest. In this Appendix, we discuss ways in which frequencies can be chosen so that $\mathbf{F}$ satisfies $\mathbf{F}^H\mathbf{F} = \mathbf{I}_N$ or $\mathbf{F}\mathbf{F}^H = \mathbf{I}_M$. As we will see, this amounts to building $\mathbf{F}$ from appropriately selected rows or columns of a DTFT matrix of appropriate dimensions.

Consider the set of uniformly spaced frequencies

$$\Omega := \left\{ \frac{2\pi k}{LN} : k = 0, 1, \ldots, LN-1 \right\} \tag{A.38}$$

which is completely specified by any nonnegative integer N and rational L such that LN is an integer greater than 1. Assume we are interested in an integer number $M \leq LN$ of these frequencies, specified by the finite sequence

$$\Gamma := \{\gamma_0, \gamma_1, \ldots, \gamma_{M-1}\} \tag{A.39}$$

where $\frac{2\pi\gamma_m}{LN} \in \Omega$. (Note that the entries of Γ do not necessarily correspond to adjacent frequencies.) We can now construct a linear operator that accepts a time domain vector and returns the DTFT of that vector evaluated only at the frequencies of interest. This operator is given by $\mathbf{F} \in \mathbb{C}^{M\times N}$ with

$$[\mathbf{F}]_{m,n} := \frac{1}{\eta}\exp\left(-j\frac{2\pi\Gamma(m)n}{LN}\right) \tag{A.40}$$

where the normalization constant η is determined by the relationship between L, N, and

M. (More on this later.) The frequency domain representation of a vector $\mathbf{x} \in \mathbb{C}^N$ is defined to be $\widehat{\mathbf{x}} := \mathbf{F}\mathbf{x}$ with

$$\widehat{x}_m := \frac{1}{\eta} \sum_{n=0}^{N_\star - 1} x_n \exp\left(-j \frac{2\pi \Gamma(m) n}{LN}\right) \tag{A.41}$$

where $N_\star = \min\{N, LN\}$. Similarly, the inverse transform of $\tilde{\mathbf{x}} \in \mathbb{C}^M$ is defined by $\mathbf{F}^H \tilde{\mathbf{x}}$ with

$$\left[\mathbf{F}^H \tilde{\mathbf{x}}\right]_n := \frac{1}{\eta} \sum_{m=0}^{M-1} \tilde{x}_m \exp\left(j \frac{2\pi \Gamma(m) n}{LN}\right) \tag{A.42}$$

Note that the unitary DTFT matrix can be constructed by letting $L = 1$, $N = M$, and $\eta = N$.

In practice, the length (N) of the time domain signal is fixed, and the FTM is specified at number of different frequencies in $[0, 2\pi)$. We can choose L so that the frequencies of interest in $[0, 2\pi)$ are arbitrarily close to points in Ω. When $L \geq 1$, the operator $\mathbf{F}$ can then be seen as appending $LN - N$ zeros to the input vector, computing the LN-point DTFT of the zero-padded vector, and then selecting the M coefficients corresponding to the frequencies of interest. When $L \leq 1$, $\mathbf{F}$ effectively truncates the last $N - LN$ entries of the input vector, computes the LN-point DTFT of the truncated vector, and select the M coefficients of interest. In either case L may have to be chosen so that Ω contains frequencies not of interest. As we will see, not every subset of Ω will result in an $\mathbf{F}$ that satisfies the requirements discussed in Sec. 2.5. So we choose the $\Gamma \subseteq \Omega$ that contains the fewest superfluous frequencies, and we construct $\mathbf{F}$ according to A.41. This framework for computing DTFT samples provides great flexibility in application. In the following sections, we discuss how to choose Γ so that $\mathbf{F}$ is valid.

A.2.2 Γ when $N \leq M$

The condition $\mathbf{F}^H\mathbf{F} = \mathbf{I}_N$ holds iff the inner product of any two columns of $\mathbf{F}$ satisfies

$$\frac{1}{\eta^2}\sum_{m=0}^{M-1}\exp\left(-j\frac{2\pi\Gamma(m)\Delta_n}{LN}\right) = \begin{cases} 1 & \Delta_n = 0 \\ 0 & \text{otherwise} \end{cases} \tag{A.43}$$

for all $\Delta_n \in \{0, 1, \ldots, N-1\}$ (i.e., orthonormal columns). This condition will hold for all Δ_n when $\eta = \sqrt{M}$, M is a divisor of LN, and Γ is of the form

$$\Gamma(m) = m_0 + \frac{LN}{M}m \tag{A.44}$$

where m_0 is an integer [140, pp. 356]. The definition of Γ in (A.39) implies $\max\{\Gamma\} = \Gamma(M-1) \leq LN - 1$. Thus, m_0 must also satisfy

$$m_0 \leq \frac{LN}{M} - 1 \tag{A.45}$$

Therefore, Γ constructed according to (A.44) and (A.45) will satisfy $\mathbf{F}^H\mathbf{F} = \mathbf{I}_N$.

A.2.3 Γ when $N \geq M$

The condition $\mathbf{F}\mathbf{F}^H = \mathbf{I}_M$ holds iff the inner product of any two rows of $\mathbf{F}$ satisfies

$$\frac{1}{\eta^2}\sum_{n=0}^{N_\star-1}\exp\left(-j\frac{2\pi\Delta_\Gamma n}{LN}\right) = \begin{cases} 1 & \Delta_\Gamma = 0 \\ 0 & \text{otherwise} \end{cases} \tag{A.46}$$

for all Δ_Γ, where Δ_Γ can equal the difference between any two elements in Γ (i.e., orthonormal rows). The η and Γ that satisfy (A.46) depend on the value of L. When $L \geq 1$, we have $N_\star = N$, and condition (A.46) will hold for all Δ_Γ if $\eta = \sqrt{N}$ and Γ is of the

form

$$\Gamma(m) = m_0 + ALm \tag{A.47}$$

where m_0 and A are integers. The condition $\Gamma(M-1) \leq LN-1$ implies A and m_0 must also satisfy the relationship

$$A \leq \frac{LN-1-m_0}{L(M-1)} \tag{A.48}$$

When $L \leq 1$, we have $N_\star = LN$, and condition (A.46) will hold for all Δ_Γ if $\eta = \sqrt{LN}$ and Γ is of the form

$$\Gamma(m) = m_0 + Am \tag{A.49}$$

The condition $\Gamma(M-1) \leq LN-1$ implies A and m_0 must also satisfy the relationship

$$A \leq \frac{LN-1-m_0}{(M-1)} \tag{A.50}$$

Therefore, Γ constructed according to (A.47) and (A.48) when $L \geq 1$, or according to (A.49) and (A.50) when $L \leq 1$, will satisfy $\mathbf{F}\mathbf{F}^H = \mathbf{I}_M$.

A.3 Derivation of Gradients

A.3.1 Preliminaries

Recall that each entry of the time vector $\mathbf{a}$ is a function of the phase vector $\boldsymbol{\phi}$ whereby $a_n = p_n e^{j\phi_n}$. For notational convenience, we shall suppress the dependence of $\mathbf{a}$ and $\widehat{\mathbf{a}}$ on $\boldsymbol{\phi}$. Clearly,

$$\frac{\partial a_n}{\partial \phi_s} = \begin{cases} j a_s & s = n \\ 0 & \text{otherwise} \end{cases} \tag{A.51}$$

and

$$\frac{\partial a_n^*}{\partial \phi_s} = \left[\frac{\partial a_n}{\partial \phi_s}\right]^* \tag{A.52}$$

Defining the LN^{th} root of unity to be

$$W_{LN} := e^{-j\frac{2\pi}{LN}} \tag{A.53}$$

we find that

$$\frac{\partial \widehat{a}_m}{\partial \phi_s} = \frac{\partial}{\partial \phi_s}\frac{1}{\eta}\sum_{n=0}^{N_\star - 1} a_n W_{LN}^{n\Gamma(m)} \tag{A.54}$$

$$= \begin{cases} \frac{j}{\eta} a_s W_{LN}^{s\Gamma(m)} & s < N_\star \\ 0 & \text{otherwise} \end{cases} \tag{A.55}$$

which makes use of the definition of $\widehat{a}_m$ given in (A.41). Similarly, we can show

$$\frac{\partial \widehat{a}_m^*}{\partial \phi_s} = \left[\frac{\partial \widehat{a}_m}{\partial \phi_s}\right]^* \tag{A.56}$$

When $\widehat{a}_m \neq 0$, we have

$$\frac{\partial}{\partial \phi_s} |\widehat{a}_m| = \frac{\partial}{\partial \phi_s} (\widehat{a}_m^* \widehat{a}_m)^{\frac{1}{2}} \tag{A.57}$$

$$= \mathrm{Re}\left\{ \frac{\widehat{a}_m^*}{|\widehat{a}_m|} \frac{\partial \widehat{a}_m}{\partial \phi_s} \right\} \tag{A.58}$$

$$= -\mathrm{Im}\left\{ \frac{\widehat{a}_m^*}{|\widehat{a}_m|} \frac{a_s}{\eta} W_{LN}^{s\Gamma(m)} \right\} \tag{A.59}$$

Note that $x/|x| = e^{j\phi}$ for any $x \in \mathbb{C}$ of the form $x = \alpha e^{j\phi}$ with $\alpha > 0$. Thus, $x/|x| \to e^{j\phi}$ as $\alpha \to 0$ from the right. As such, we shall let

$$\widehat{a}_m^*/|\widehat{a}_m| = e^{-j\angle \widehat{a}_m} \tag{A.60}$$

when $|\widehat{a}_m| = 0$. The partial of $|\widehat{a}_m|$ is then

$$\frac{\partial}{\partial \phi_s} |\widehat{a}_m| = -\mathrm{Im}\left\{ \frac{a_s}{\eta} e^{-j\angle \widehat{a}_m} W_{LN}^{s\Gamma(m)} \right\} \tag{A.61}$$

A.3.2 Spectrum Shaping Problem

Note that the cost function J defined in (2.23) can be written as

$$J(\boldsymbol{\phi}) = ||\mathbf{q}||^2 + \sum_{m=0}^{M-1} |\widehat{a}_m|^2 - 2q_m |\widehat{a}_m| \tag{A.62}$$

As such,

$$\frac{\partial J}{\partial \phi_s}(\boldsymbol{\phi}) = \sum_{m=0}^{M-1} \frac{\partial}{\partial \phi_s} |\widehat{a}_m|^2 - 2q_m \frac{\partial}{\partial \phi_s} |\widehat{a}_m| \tag{A.63}$$

$$= 2\mathrm{Im}\left\{ \frac{a_s}{\eta} \sum_{m=0}^{M-1} \left(q_m e^{-j\angle \widehat{a}_m} - \widehat{a}_m^* \right) W_{LN}^{s\Gamma(m)} \right\}$$

which implies

$$\nabla_{\boldsymbol{\phi}}^T J = 2\mathrm{Im}\left\{\mathbf{a} \odot \left[\mathbf{F}^H \left(P_{\widetilde{B}}(\mathbf{Fa}) - \mathbf{Fa}\right)\right]^*\right\} \tag{A.64}$$

Note that $\mathbf{F}^H\mathbf{F} = \mathbf{I}_N$ when $N \leq M$. In this case, the gradient becomes

$$\nabla_{\boldsymbol{\phi}}^T J = 2\mathrm{Im}\left\{\mathbf{a} \odot \left[\mathbf{F}^H P_{\widetilde{B}}(\mathbf{Fa})\right]^*\right\} \tag{A.65}$$

A.3.3 PSD Synthesis Problem

The cost J_2 defined in (2.30) can be written as

$$J_2(\boldsymbol{\phi}) = \left|\left|\mathbf{q}^2\right|\right|^2 + \sum_{m=1}^{M-1} \left|\widehat{a}_m\right|^4 - 2q_m^2 \left|\widehat{a}_m\right|^2 \tag{A.66}$$

As such,

$$\frac{\partial J_2}{\partial \phi_s}(\boldsymbol{\phi}) = \sum_{m=0}^{M-1} \frac{\partial}{\partial \phi_s} \left|\widehat{a}_m\right|^4 - 2q_m^2 \frac{\partial}{\partial \phi_s} \left|\widehat{a}_m\right|^2 \tag{A.67}$$

$$= 4\sum_{m=0}^{M-1} \left(\left|\widehat{a}_m\right|^3 - q_m^2 \left|\widehat{a}_m\right|\right) \frac{\partial}{\partial \phi_s} \left|\widehat{a}_m\right| \tag{A.68}$$

$$= 4\mathrm{Im}\left\{\frac{a_s}{\eta} \sum_{m=0}^{M-1} \left(q_m^2 - \left|\widehat{a}_m\right|^2\right) \widehat{a}_m^* W_{LN}^{s\Gamma(m)}\right\}$$

Which implies

$$\nabla_{\boldsymbol{\phi}}^T J_2 = 2\mathrm{Im}\left\{\mathbf{a} \odot \left[\mathbf{F}^H \left(\left(\mathbf{q}^2 - \left|\mathbf{Fa}\right|^2\right) \odot \mathbf{Fa}\right)\right]^*\right\} \tag{A.69}$$

Appendix B

Appendix for Chapter 3

B.1 SINR Gradients

B.1.1 SINR for Joint Design

For the joint signal/filter design, SINR (3.36) is a function of both the signal phase vector $\boldsymbol{\phi} \in \mathbb{R}^N$ and the filter $\mathbf{h} \in \mathbb{C}^N$. Consider the partial of (3.36) with respect to the p^{th} element of the signal phase vector ϕ. This is given by

$$\frac{\partial}{\partial \phi_p} \text{SINR}\left(\mathbf{s}, \mathbf{h}\right) = \frac{\frac{\partial}{\partial \phi_p}\left|\mathbf{h}^H \mathbf{s}\right|^2}{\mathbf{h}^H \mathbf{K} \mathbf{h}}. \tag{B.1}$$

To evaluate the numerator, note that

$$\frac{\partial s_n}{\partial \phi_p} = \begin{cases} j s_p & p = n \\ 0 & \text{otherwise} \end{cases} \tag{B.2}$$

and

$$\frac{\partial s_n^*}{\partial \phi_p} = \left[\frac{\partial s_n}{\partial \phi_p}\right]^* \tag{B.3}$$

Then,

$$\frac{\partial}{\partial \phi_p}\left|\mathbf{h}^H\mathbf{s}\right|^2 = \frac{\partial}{\partial \phi_p}\left[(\mathbf{h}^H\mathbf{s})(\mathbf{h}^H\mathbf{s})^*\right] \tag{B.4}$$

$$= 2\mathrm{Re}\left\{(\mathbf{h}^H\mathbf{s})\frac{\partial}{\partial \phi_p}\mathbf{s}^H\mathbf{h}\right\} \tag{B.5}$$

$$= 2\mathrm{Im}\left\{(\mathbf{h}^H\mathbf{s})h_p s_p^*\right\} \tag{B.6}$$

The gradient of (3.36) with respect to the signal phase vector can then be found by substituting (B.6) into (B.1), and then stacking all the partials into a vector. This results in

$$\nabla_{\boldsymbol{\phi}}^T \mathrm{SINR}\left(\mathbf{s},\mathbf{h}\right) = \frac{2\mathrm{Im}\left\{(\mathbf{h}^H\mathbf{s})(\mathbf{h}\odot\mathbf{s}^*)\right\}}{\mathbf{h}^H\mathbf{K}\mathbf{h}}. \tag{B.7}$$

Next, consider the partial of (3.36) with respect to the p^{th} element of the filter $\mathbf{h}$. Applying the quotient rule yields

$$\frac{\partial}{\partial h_p}\mathrm{SINR}\left(\mathbf{s},\mathbf{h}\right) = \frac{\left(\mathbf{h}^H\mathbf{K}\mathbf{h}\right)\frac{\partial}{\partial h_p}\left|\mathbf{h}^H\mathbf{s}\right|^2 - \left|\mathbf{h}^H\mathbf{s}\right|^2\frac{\partial}{\partial h_p}\mathbf{h}^H\mathbf{K}\mathbf{h}}{\left(\mathbf{h}^H\mathbf{K}\mathbf{h}\right)^2} \tag{B.8}$$

Evaluating the partials in the numerator, we find

$$\frac{\partial}{\partial h_p}\left|\mathbf{h}^H\mathbf{s}\right|^2 = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} s_m h_m^* s_n^* \frac{\partial h_n}{\partial h_p} \tag{B.9}$$

$$= s_p^*\sum_{m=0}^{N-1} s_m h_m^* \tag{B.10}$$

$$= s_p^*(\mathbf{h}^H\mathbf{s}). \tag{B.11}$$

and

$$\frac{\partial}{\partial h_p}\mathbf{h}^H\mathbf{K}\mathbf{h} = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} h_n^*[\mathbf{K}]_{n,m}\frac{\partial h_m}{\partial h_p} \tag{B.12}$$

$$= \sum_{n=0}^{N-1} h_n^*[\mathbf{K}]_{n,p} \tag{B.13}$$

$$= \mathbf{h}^H\mathbf{k}_p. \tag{B.14}$$

where $\mathbf{k}_p$ denotes the p^{th} column of $\mathbf{K}$. Substituting (B.11) and (B.14) into (B.8), and then stacking all the partials into a vector yields the gradient of the SINR with respect to the filter. This is given by

$$\nabla_{\mathbf{h}}\text{SINR}\left(\mathbf{s},\mathbf{h}\right) = \alpha\mathbf{s}^H - \left|\alpha\right|^2\mathbf{h}^H\mathbf{K} \tag{B.15}$$

where $\alpha := \mathbf{h}^H\mathbf{s}/\mathbf{h}^H\mathbf{K}\mathbf{h}$.

B.1.2 SINR for WF and MF Designs

For both the whitening filter (i.e., Neyman-Pearson) and matched filter designs, the SINR is a function only of the signal phase vector $\boldsymbol{\phi}$. Let $\mathbf{K} := [\mathbf{k}_1 \cdots \mathbf{k}_N]$ be a Hermitian matrix. Then,

$$\frac{\partial}{\partial\phi_p}\mathbf{s}^H\mathbf{K}\mathbf{s} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1}[\mathbf{K}]_{n,m}\frac{\partial}{\partial\phi_p}(s_n^*s_m) \tag{B.16}$$

$$= (js_m)\sum_{n=0}^{N-1}[\mathbf{K}]_{n,p}s_n^* + (-js_p^*)\sum_{m=0}^{N-1}[\mathbf{K}]_{p,m}s_m \tag{B.17}$$

$$= 2\text{Re}\left\{(-js_p^*)\mathbf{k}_p^H\mathbf{s}\right\} \tag{B.18}$$

$$= 2\text{Im}\left\{s_p^*\mathbf{k}_p^H\mathbf{s}\right\} \tag{B.19}$$

Stacking the partials into a vector yields the gradient of the matched filter design objective function in (3.47). This gradient is given by

$$\nabla_{\boldsymbol{\phi}}^{T}\mathbf{s}^{H}\mathbf{K}\mathbf{s} = 2\mathrm{Im}\left\{\mathbf{s}^{*}\odot(\mathbf{K}\mathbf{s})\right\}. \tag{B.20}$$

Similarly, the gradient of the whitening filter design objective function in (3.46) is

$$\nabla_{\boldsymbol{\phi}}^{T}\mathbf{s}^{H}\mathbf{K}^{-1}\mathbf{s} = 2\mathrm{Im}\left\{\mathbf{s}^{*}\odot(\mathbf{K}^{-1}\mathbf{s})\right\}. \tag{B.21}$$

B.2 Normalized XCS Jacobians

B.2.1 Normalized XCS for Joint Design

For the joint signal/filter design problem, the normalized XCS is a function of both the signal phase vector $\boldsymbol{\phi}\in\mathbb{R}^{N}$ and the filter $\mathbf{h}\in\mathbb{C}^{N}$. Let $k\geq 0$, and consider the partial of the squared magnitude of the normalized XCS at lag k with respect to the p^{th} element of the signal phase. By the quotient rule we have

$$\frac{\partial}{\partial\phi_p}\left|\tilde{R}_k(\mathbf{s},\mathbf{h})\right|^2 = \frac{\left|\mathbf{h}^H\mathbf{s}\right|^2\frac{\partial}{\partial\phi_p}\left|R_k(\mathbf{s},\mathbf{h})\right|^2 - \left|R_k(\mathbf{s},\mathbf{h})\right|^2\frac{\partial}{\partial\phi_p}\left|\mathbf{h}^H\mathbf{s}\right|^2}{\left|\mathbf{h}^H\mathbf{s}\right|^4} \tag{B.22}$$

In order to evaluate the numerator, we first note that

$$\frac{\partial}{\partial\phi_p}R_k(\mathbf{s},\mathbf{h}) = \sum_{n=0}^{N-k-1} h_n^{*}\frac{\partial}{\partial\phi_p}s_{n+k} \tag{B.23}$$

$$= \begin{cases} js_p h_{p-k}^{*} & k\leq p \\ 0 & \text{otherwise} \end{cases} \tag{B.24}$$

$$= js_p h_{p-k}^{*}u(p-k) \tag{B.25}$$

where u denotes the unit step function with $u(n) = 1$ for $n \geq 0$ and $u(n) = 0$ for $n < 0$. Also,

$$\frac{\partial}{\partial \phi_p} R_k^*(\mathbf{s}, \mathbf{h}) = \left[\frac{\partial}{\partial \phi_p} R_k(\mathbf{s}, \mathbf{h}) \right]^* . \tag{B.26}$$

Using this conjugate symmetry, one can show

$$\frac{\partial}{\partial \phi_p} \left| R_k(\mathbf{s}, \mathbf{h}) \right|^2 = 2\mathrm{Im} \left\{ R_k(\mathbf{s}, \mathbf{h}) s_p^* h_{p-k} \right\} u(p-k) \tag{B.27}$$

Substituting (B.6) and (B.27) into (B.1), and then simplifying, yields

$$\frac{\partial}{\partial \phi_p} \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 = 2\mathrm{Im} \left\{ \frac{s_p^* h_{p-k}}{|\mathbf{h}^H \mathbf{s}|} \tilde{R}_k(\mathbf{s}, \mathbf{h}) u(p-k) - \frac{s_p^* h_p}{(\mathbf{h}^H \mathbf{s})^*} \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 \right\} . \tag{B.28}$$

Stacking the partials into gradient vectors, and then combining the gradients to form the Jacobian, yields

$$J_{\boldsymbol{\phi}}^T \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 = 2\mathrm{Im} \left\{ \frac{\mathbf{R} \odot \mathbf{H}}{|\mathbf{h}^H \mathbf{s}|} - \frac{(\mathbf{h} \odot \mathbf{s}^*) \otimes \mathbf{r}^T}{(\mathbf{h}^H \mathbf{s})^*} \right\} \tag{B.29}$$

where

$$[\mathbf{R}]_{p,k} := \tilde{R}_k(\mathbf{s}, \mathbf{h}) \tag{B.30}$$

$$[\mathbf{H}]_{p,k} := \begin{cases} s_p^* h_{p-k} & k \leq p \\ 0 & \text{otherwise} \end{cases} \tag{B.31}$$

$$[\mathbf{r}]_k := \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 \tag{B.32}$$

Similarly, the Jacobian of the squared magnitude of the normalized XCS evaluated at the negative lags is given by

$$J_{\boldsymbol{\phi}}^{T}\left|\tilde{R}_{-k}(\mathbf{s},\mathbf{h})\right|^{2}=2\mathrm{Im}\left\{\frac{\widehat{\mathbf{R}}\odot\widehat{\mathbf{H}}}{|\mathbf{h}^{H}\mathbf{s}|}-\frac{(\mathbf{h}\odot\mathbf{s}^{*})\otimes\widehat{\mathbf{r}}^{T}}{(\mathbf{h}^{H}\mathbf{s})^{*}}\right\} \tag{B.33}$$

where

$$[\widehat{\mathbf{R}}]_{p,k}:=\tilde{R}_{-k}(\mathbf{s},\mathbf{h}) \tag{B.34}$$

$$[\widehat{\mathbf{H}}]_{p,k}:=\begin{cases}s_{p}^{*}h_{p+k} & p<N-k\\ 0 & \text{otherwise}\end{cases} \tag{B.35}$$

$$[\widehat{\mathbf{r}}]_{k}:=\left|\tilde{R}_{-k}(\mathbf{s},\mathbf{h})\right|^{2} \tag{B.36}$$

Consider the partial of the squared magnitude of normalized XCS at lag k with respect to the p^{th} filter element. By the quotient rule we have

$$\frac{\partial}{\partial h_{p}}\left|\tilde{R}_{k}(\mathbf{s},\mathbf{h})\right|^{2}=\frac{\left|\mathbf{h}^{H}\mathbf{s}\right|^{2}\frac{\partial}{\partial h_{p}}\left|R_{k}(\mathbf{s},\mathbf{h})\right|^{2}-\left|R_{k}(\mathbf{s},\mathbf{h})\right|^{2}\frac{\partial}{\partial h_{p}}\left|\mathbf{h}^{H}\mathbf{s}\right|^{2}}{\left|\mathbf{h}^{H}\mathbf{s}\right|^{4}} \tag{B.37}$$

In order to evaluate the numerator, we first note that

$$\frac{\partial}{\partial h_{p}}R_{k}(\mathbf{s},\mathbf{h})=\sum_{n=0}^{N-k-1}s_{n+k}\frac{\partial h_{n}^{*}}{\partial h_{p}}=0 \tag{B.38}$$

and (*continued overleaf*)

$$\frac{\partial}{\partial h_p} R_k^*(\mathbf{s}, \mathbf{h}) = \sum_{n=0}^{N-k-1} s_{n+k}^* \frac{\partial h_n}{\partial h_p} \tag{B.39}$$

$$= \begin{cases} s_{p+k}^* & p < N-k \\ 0 & \text{otherwise} \end{cases} \tag{B.40}$$

$$= s_{p+k}^* u(N-k-p) \tag{B.41}$$

Thus,

$$\frac{\partial}{\partial h_p} \left| R_k(\mathbf{s}, \mathbf{h}) \right|^2 = \frac{\partial}{\partial h_p} \left(R_k^*(\mathbf{s}, \mathbf{h}) R_k(\mathbf{s}, \mathbf{h}) \right) \tag{B.42}$$

$$= R_k(\mathbf{s}, \mathbf{h}) s_{p+k}^* u(N-k-p) \tag{B.43}$$

Substituting (B.11) and (B.43) into (B.37), and then simplifying, yields

$$\frac{\partial}{\partial h_p} \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 = \frac{s_{p+k}^*}{|\mathbf{h}^H \mathbf{s}|} \tilde{R}_k(\mathbf{s}, \mathbf{h}) u(N-k-p) - \left(\frac{s_p}{\mathbf{h}^H \mathbf{s}} \right)^* \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 \tag{B.44}$$

By stacking the partials into gradient vectors, and then combining the gradients to form the Jacobian, the corresponding Jacobian can be computed as

$$J_{\mathbf{h}}^T \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 = \frac{\mathbf{R} \odot \mathbf{S}}{|\mathbf{h}^H \mathbf{s}|} - \frac{\mathbf{s}^* \otimes \mathbf{r}^T}{(\mathbf{h}^H \mathbf{s})^*} \tag{B.45}$$

where $\mathbf{R}$ and $\mathbf{r}$ are defined in (B.30) and (B.32), respectively, and

$$[\mathbf{S}]_{p,k} := \begin{cases} s_{p+k}^* & p < N-k \\ 0 & \text{otherwise} \end{cases}. \tag{B.46}$$

In a similar way, it can be shown that

$$J_{\mathbf{h}}^T \left| \tilde{R}_{-k}(\mathbf{s}, \mathbf{h}) \right|^2 = \frac{\widehat{\mathbf{R}} \odot \widehat{\mathbf{S}}}{|\mathbf{h}^H \mathbf{s}|} - \frac{\mathbf{s}^* \otimes \widehat{\mathbf{r}}^T}{(\mathbf{h}^H \mathbf{s})^*} \tag{B.47}$$

where $\widehat{\mathbf{R}}$ and $\widehat{\mathbf{r}}$ are defined in (B.34) and (B.36), respectively, and

$$[\widehat{\mathbf{S}}]_{p,k} := \begin{cases} s_{p-k}^* & k \leq p \\ 0 & \text{otherwise} \end{cases}. \tag{B.48}$$

B.2.2 Normalized XCS for WF Design

For the whitening filter optimization problem in (3.46), the normalized XCS is a function of the signal phase vector $\boldsymbol{\phi} \in \mathbb{R}^N$ only. Let $k > 0$, and consider the partial of the squared magnitude of the normalized XCS at lag k with respect to the p^{th} element of the phase vector. For simplicity, replace $\mathbf{K}^{-1}$ with an arbitrary positive definite matrix $\mathbf{F}$. Then, by the quotient rule

$$\frac{\partial}{\partial \phi_p} \left| \tilde{R}_k(\mathbf{s}, \mathbf{F}\mathbf{s}) \right|^2 = \frac{\left|\mathbf{s}^H \mathbf{F} \mathbf{s}\right|^2 \frac{\partial}{\partial \phi_p} \left|R_k(\mathbf{s}, \mathbf{F}\mathbf{s})\right|^2 - \left|R_k(\mathbf{s}, \mathbf{F}\mathbf{s})\right|^2 \frac{\partial}{\partial \phi_p} \left|\mathbf{s}^H \mathbf{F} \mathbf{s}\right|^2}{\left|\mathbf{s}^H \mathbf{F} \mathbf{s}\right|^4} \tag{B.49}$$

The partial of the XCS with respect to the p^{th} element of the phase vector $\boldsymbol{\phi}$ is given by

$$\begin{aligned}
\frac{\partial}{\partial \phi_p} R_k(\mathbf{s}, \mathbf{F}\mathbf{s}) &= \frac{\partial}{\partial \phi_p} \sum_{n=0}^{N-k-1} s_{n+k} [\mathbf{F}\mathbf{s}]_n^* && \text{(B.50)} \\
&= \frac{\partial}{\partial \phi_p} \sum_{n=0}^{N-k-1} s_{n+k} \left[\sum_{m=0}^{N-1} [\mathbf{F}]_{n,m} s_m \right]^* && \text{(B.51)} \\
&= \sum_{n=0}^{N-k-1} \sum_{m=0}^{N-1} [\mathbf{F}]_{n,m}^* \frac{\partial}{\partial \phi_p} (s_m^* s_{n+k}) && \text{(B.52)} \\
&= \sum_{n=0}^{N-k-1} [\mathbf{F}]_{n,p}^* s_{n+k} (-j s_p^*) + \sum_{m=0}^{N-1} [\mathbf{F}]_{p-k,m}^* s_m^* (j s_p) u(p-k) && \text{(B.53)} \\
&= (-j s_p^*) R_k(\mathbf{s}, \mathbf{f}_p) + (j s_p) [\mathbf{F}\mathbf{s}]_{p-k}^* u(p-k) && \text{(B.54)}
\end{aligned}$$

where $\mathbf{f}_p \in \mathbb{C}^N$ is the p^{th} column of $\mathbf{F}$. In a similar way, one can show that

$$\frac{\partial}{\partial \phi_p} R_k^*(\mathbf{s}, \mathbf{Fs}) = \left[\frac{\partial}{\partial \phi_p} R_k(\mathbf{s}, \mathbf{Fs}) \right]^*. \tag{B.55}$$

Using this conjugate symmetry, we have

$$\frac{\partial}{\partial \phi_p} \left| R_k(\mathbf{s}, \mathbf{Fs}) \right|^2 = 2\mathrm{Im} \left\{ R_k^*(\mathbf{s}, \mathbf{Fs}) \left(s_p^* R_k(\mathbf{s}, \mathbf{f}_p) - s_p [\mathbf{Fs}]_{p-k}^* u(p-k) \right) \right\} \tag{B.56}$$

Since, $\mathbf{F}$ is positive definite, then

$$\frac{\partial}{\partial \phi_p} \left(\mathbf{s}^H \mathbf{Fs} \right) = \left[\frac{\partial}{\partial \phi_p} \mathbf{s}^H \mathbf{Fs} \right]^*. \tag{B.57}$$

This result, along with (B.19), implies

$$\frac{\partial}{\partial \phi_p} \left| \mathbf{s}^H \mathbf{Fs} \right|^2 = 2\mathrm{Re} \left\{ \left(\mathbf{s}^H \mathbf{Fs} \right) \frac{\partial}{\partial \phi_p} \mathbf{s}^H \mathbf{Fs} \right\} \tag{B.58}$$

$$= 4 \left(\mathbf{s}^H \mathbf{Fs} \right) \mathrm{Im} \left\{ s_p^* \mathbf{f}_p^H \mathbf{s} \right\} \tag{B.59}$$

Substituting (B.56) and (B.59) into (B.49), applying the substitution $\mathbf{h} = \mathbf{Fs}$, recognizing that $\mathbf{f}_p^H \mathbf{s} = h_p$, and simplifying, yields

$$\frac{\partial}{\partial \phi_p} \left| \tilde{R}_k(\mathbf{s}, \mathbf{Fs}) \right|^2 = \frac{2\mathrm{Im} \left\{ \tilde{R}_k^*(\mathbf{s}, \mathbf{h}) \left(s_p^* R_k(\mathbf{s}, \mathbf{f}_p) - s_p h_{p-k}^* u(p-k) \right) - 2 \left| \tilde{R}_k(\mathbf{s}, \mathbf{h}) \right|^2 s_p^* h_p \right\}}{\mathbf{h}^H \mathbf{s}}. \tag{B.60}$$

The corresponding Jacobian can be computed as

$$J_{\boldsymbol{\phi}}^T \left| \tilde{R}_k(\mathbf{s}, \mathbf{Fs}) \right|^2 = \frac{2\mathrm{Im} \left\{ \mathbf{R}^* \left(\mathbf{T} - \mathbf{H}^* \right) - 2 \left(\mathbf{h} \odot \mathbf{s}^* \right) \otimes \mathbf{r}^T \right\}}{\mathbf{h}^H \mathbf{s}} \tag{B.61}$$

where $\mathbf{R}$, $\mathbf{H}$, and $\mathbf{r}$ are defined in (B.30), (B.31), and (B.32), respectively, and

$$[\mathbf{T}]_{p,k} := s_p^* R_k(\mathbf{s}, \mathbf{f}_p). \tag{B.62}$$

Similarly, it can be shown that

$$J_{\boldsymbol{\phi}}^T \left| \tilde{R}_{-k}(\mathbf{s}, \mathbf{F}\mathbf{s}) \right|^2 = \frac{2\mathrm{Im}\left\{ \widehat{\mathbf{R}}^* \left(\widehat{\mathbf{T}} - \widehat{\mathbf{H}}^* \right) - 2\left(\mathbf{h} \odot \mathbf{s}^*\right) \otimes \widehat{\mathbf{r}}^T \right\}}{\mathbf{h}^H \mathbf{s}} \tag{B.63}$$

where $\widehat{\mathbf{R}}$, $\widehat{\mathbf{H}}$, and $\widehat{\mathbf{r}}$ are defined in (B.34), (B.35), and (B.36), respectively, and

$$[\widehat{\mathbf{T}}]_{p,k} := s_p^* R_{-k}(\mathbf{s}, \mathbf{f}_p). \tag{B.64}$$

B.2.3 Normalized XCS for MF Design

For the matched filter optimization problem in (3.47), the normalized ACS is a function of the signal phase vector $\boldsymbol{\phi} \in \mathbb{R}^N$ only. Furthermore, we shall assume $||\mathbf{s}|| = 1$ so that $\tilde{R}_k = R_k$. Let $k > 0$, and consider the partial of the squared magnitude of the ACS at lag k with respect to the p^{th} element of the phase vector. This is given by

$$\frac{\partial}{\partial \phi_p} R_k(\mathbf{s}, \mathbf{s}) = \sum_{n=0}^{N-k-1} \frac{\partial}{\partial \phi_p} (s_{n+k} s_n^*) \tag{B.65}$$

$$= (j s_p)^* s_{p+k} u(N-k-p) + (j s_p) s_{p-k}^* u(p-k) \tag{B.66}$$

Similarly, one can show that

$$\frac{\partial}{\partial \phi_p} R_k^*(\mathbf{s}, \mathbf{s}) = [R_k^*(\mathbf{s}, \mathbf{s})]^* \tag{B.67}$$

Using this conjugate symmetry, we find

$$\frac{\partial}{\partial \phi_p} \left| R_k(\mathbf{s}, \mathbf{s}) \right|^2 = 2\mathrm{Im} \left\{ R_k(\mathbf{s}, \mathbf{s}) \left(s_p^* s_{p+k} u(N-k-p) - s_p s_{p-k}^* u(p-k) \right) \right\} \tag{B.68}$$

The corresponding Jacobian can be computed as

$$J_{\boldsymbol{\phi}}^T \left| R_k(\mathbf{s}, \mathbf{s}) \right|^2 = 2\mathrm{Im} \left\{ \mathbf{R}(\widehat{\mathbf{H}} - \mathbf{H}^*) \right\} \tag{B.69}$$

where by recognizing that $\mathbf{h} = \mathbf{s}$, we can define $\mathbf{R}$, $\mathbf{H}$, and $\widehat{\mathbf{H}}$ by (B.30), (B.31) and (B.35), respectively. Note that due to ACS symmetry, only the positive lags need to be constrained.

Appendix C

Appendix for Chapter 4

C.1 Gradients for Magnitude and Phase Design

C.1.1 SINR Gradient

In order to find the partial of (4.1) with respect to the p^{th} entry in the waveform $\mathbf{s}$, we apply the quotient rule to find

$$\frac{\partial}{\partial s_p}F(\mathbf{s}) = \frac{\mathbf{s}^H\mathbf{K}\mathbf{s}\left(\frac{\partial}{\partial s_p}\left|\mathbf{s}^H\mathbf{s}\right|^2\right) - \left|\mathbf{s}^H\mathbf{s}\right|^2\left(\frac{\partial}{\partial s_p}\mathbf{s}^H\mathbf{K}\mathbf{s}\right)}{\left(\mathbf{s}^H\mathbf{K}\mathbf{s}\right)^2}. \tag{C.1}$$

Evaluating the numerator, we have

$$\begin{aligned}
\frac{\partial}{\partial s_p}\left|\mathbf{s}^H\mathbf{s}\right|^2 &= \frac{\partial}{\partial s_p}(\mathbf{s}^H\mathbf{s})^*(\mathbf{s}^H\mathbf{s}) && \text{(C.2)}\\
&= (\mathbf{s}^H\mathbf{s})^*\frac{\partial}{\partial s_p}(\mathbf{s}^H\mathbf{s}) + (\mathbf{s}^H\mathbf{s})\frac{\partial}{\partial s_p}(\mathbf{s}^H\mathbf{s})^* && \text{(C.3)}\\
&= 2(\mathbf{s}^H\mathbf{s})\frac{\partial}{\partial s_p}(\mathbf{s}^H\mathbf{s}) && \text{(C.4)}\\
&= 2(\mathbf{s}^H\mathbf{s})\sum_{n=0}^{N-1}\frac{\partial}{\partial s_p}(s_n^*s_n) && \text{(C.5)}\\
&= 2\left|\mathbf{s}^H\mathbf{s}\right|s_p^* && \text{(C.6)}
\end{aligned}$$

and

$$\frac{\partial}{\partial s_p}\mathbf{s}^H\mathbf{K}\mathbf{s} = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\frac{\partial}{\partial s_p}[\mathbf{K}]_{n,m}s_n^*s_m = \mathbf{s}^H\mathbf{k}_p \tag{C.7}$$

where $\mathbf{k}_p$ is the p^{th} column of $\mathbf{K}$. Substituting (C.6) and (C.7) into (C.1), we have

$$\frac{\partial}{\partial s_p}F(\mathbf{s}) = 2\frac{\left|\mathbf{s}^H\mathbf{s}\right|}{\mathbf{s}^H\mathbf{K}\mathbf{s}}s_p^* - \frac{\left|\mathbf{s}^H\mathbf{s}\right|^2}{(\mathbf{s}^H\mathbf{K}\mathbf{s})^2}\mathbf{s}^H\mathbf{k}_p. \tag{C.8}$$

Arranging all of the partial into a vector yields the gradient

$$\nabla_{\mathbf{s}}F(\mathbf{s}) = 2\frac{\left|\mathbf{s}^H\mathbf{s}\right|}{\mathbf{s}^H\mathbf{K}\mathbf{s}}\mathbf{s}^H - \frac{\left|\mathbf{s}^H\mathbf{s}\right|^2}{(\mathbf{s}^H\mathbf{K}\mathbf{s})^2}\mathbf{s}^H\mathbf{K}. \tag{C.9}$$

C.1.2 ACS Constraint Jacobian

The partial of the normalized ACS in (4.3) with respect to the p^{th} element of the vector s can be found by applying the quotient rule, which yields

$$\frac{\partial}{\partial s_p}\left|\tilde{R}_k(\mathbf{s})\right|^2 = \frac{\partial}{\partial s_p}\frac{\left|R_k(\mathbf{s})\right|^2}{\left|\mathbf{s}^H\mathbf{s}\right|^2} \tag{C.10}$$

$$= \frac{1}{\left|\mathbf{s}^H\mathbf{s}\right|^4}\left(\left|\mathbf{s}^H\mathbf{s}\right|^2\frac{\partial}{\partial s_p}\left|R_k(\mathbf{s})\right|^2 - \left|R_k(\mathbf{s})\right|^2\frac{\partial}{\partial s_p}\left|\mathbf{s}^H\mathbf{s}\right|^2\right) \tag{C.11}$$

We note that

$$\frac{\partial}{\partial s_p}R_k(\mathbf{s}) = \sum_{n=0}^{N-k-1}\left(\frac{\partial}{\partial s_p}s_{n+k}\right)s_n^* = s_{p-k}^*u(p-k) \tag{C.12}$$

and

$$\frac{\partial}{\partial s_p}R_k^*(\mathbf{s}) = \sum_{n=0}^{N-k-1}s_{n+k}^*\left(\frac{\partial}{\partial s_p}s_n^*\right) = s_{p+k}^*u(N-k-1-p) \tag{C.13}$$

where $u : \mathbb{R} \to \mathbb{R}$ is the unit step function (i.e., $u(t)$ is equal to zero except when $t \geq 0$). We can use (C.12) and (C.13) to evaluate the first partial in the numerator of (C.11).

$$\frac{\partial}{\partial s_p} |R_k(\mathbf{s})|^2 = \frac{\partial}{\partial s_p} R_k^*(\mathbf{s}) R_k(\mathbf{s}) \tag{C.14}$$

$$= R_k^*(\mathbf{s}) \frac{\partial}{\partial s_p} R_k(\mathbf{s}) + R_k(\mathbf{s}) \frac{\partial}{\partial s_p} R_k^*(\mathbf{s}) \tag{C.15}$$

$$= R_k^*(\mathbf{s}) s_{p-k}^* u(p-k) + R_k(\mathbf{s}) s_{p+k}^* u(N-k-1-p) \tag{C.16}$$

Substituting (C.16) and (C.6) into (C.11), and simplifying, yields

$$\frac{\partial}{\partial s_p} \left| \tilde{R}_k(\mathbf{s}) \right|^2 = \frac{1}{|\mathbf{s}^H \mathbf{s}|} \left(\tilde{R}_k^*(\mathbf{s}) s_{p-k}^* u(p-k) + \tilde{R}_k(\mathbf{s}) s_{p+k}^* u(N-k-1-p) \right) - \frac{2}{|\mathbf{s}^H \mathbf{s}|} \left| \tilde{R}_k(\mathbf{s}) \right|^2 s_p^* \tag{C.17}$$

Arranging the N partials of the N functions into a matrix yields the Jacobian:

$$J_{\mathbf{s}}^T \left| \tilde{R}_k(\mathbf{s}) \right|^2 = \frac{\mathbf{M} + \mathbf{N} - 2\left(\mathbf{s}^* \otimes \mathbf{r}^T\right)}{|\mathbf{s}^H \mathbf{s}|}, \tag{C.18}$$

where the matrices $\mathbf{M}, \mathbf{N} \in \mathbb{C}^{N \times N}$ and the vector $\mathbf{r} \in \mathbb{R}_+^N$ are defined as

$$[\mathbf{M}]_{p,k} := \begin{cases} \tilde{R}_k^*(\mathbf{s}) s_{p-k}^* & k \leq p \\ 0 & \text{otherwise} \end{cases} \tag{C.19}$$

$$[\mathbf{N}]_{p,k} := \begin{cases} \tilde{R}_k(\mathbf{s}) s_{p+k}^* & p \leq N-k-1 \\ 0 & \text{otherwise} \end{cases} \tag{C.20}$$

$$[\mathbf{r}]_k := \left| \tilde{R}_k(\mathbf{s}) \right|^2. \tag{C.21}$$

C.1.3 Similarity Constraint Gradient

Consider the similarity constraint function in (4.5). Neither $\mathbf{s}_d$ nor ϵ depend on $\mathbf{s}$. Thus,

$$\nabla_{\mathbf{s}}\mathcal{S} = \nabla_{\mathbf{s}}\left[(\mathbf{s}_d - \mathbf{s})^H(\mathbf{s}_d - \mathbf{s}) - \epsilon\right] \tag{C.22}$$

$$= \nabla_{\mathbf{s}}\left[\mathbf{s}_d^H\mathbf{s}_d - \mathbf{s}_d^H\mathbf{s} - \mathbf{s}^H\mathbf{s}_d + \mathbf{s}^H\mathbf{s} - \epsilon\right] \tag{C.23}$$

$$= (\mathbf{s} - \mathbf{s}_d)^H \tag{C.24}$$

C.1.4 Modulus Constraint Jacobian

Consider the n^{th} modulus constraint function defined in (4.6). Since $\mathbf{p}$ does not depend on $\mathbf{s}$, we have

$$\frac{\partial M_n}{\partial x_p}(\mathbf{s}) = \frac{\partial}{\partial x_p}\left(s_n^* s_n - p_n\right) = \begin{cases} s_p^* & n = p \\ 0 & n \neq p \end{cases} \tag{C.25}$$

Thus, the gradient becomes

$$\nabla_{\mathbf{s}} M_n = [0\ 0\ \ldots\ 0\ s_n^*\ 0\ \ldots\ 0] \tag{C.26}$$

which is a row vector of zeros with s_n^* in the n^{th} position. The Jacobian of $\mathcal{M}$ in (4.7) is then given by

$$J_{\mathcal{M}} = \text{diag}\left\{\mathbf{s}^H\right\} \tag{C.27}$$

which is a matrix of zeros with $\mathbf{s}^H$ on the diagonal.

C.2 Gradients for Phase-Only Design

C.2.1 SINR Gradient

Following the derivation in Sec. B.1.2, the gradient of (4.9) with respect to the phase vector $\boldsymbol{\phi} \in \mathbb{R}^N$ is given by

$$\nabla_{\boldsymbol{\phi}} \widehat{F}(\boldsymbol{\phi}) = 2\mathrm{Im}\left\{\mathbf{s}^* \odot (\mathbf{K}\mathbf{s})\right\}. \tag{C.28}$$

C.2.2 ACS Constraint Jacobian

Following the derivation in Sec. B.2.3, we have

$$J_{\boldsymbol{\phi}}^T \left|R_k(\mathbf{s},\mathbf{s})\right|^2 = 2\mathrm{Im}\left\{\mathbf{R}(\widehat{\mathbf{H}} - \mathbf{H}^*)\right\} \tag{C.29}$$

where by recognizing that $\mathbf{h} = \mathbf{s}$, we can define $\mathbf{R}$, $\mathbf{H}$, and $\widehat{\mathbf{H}}$ by (B.30), (B.31) and (B.35), respectively.

C.2.3 Similarity Constraint Gradient

Let $\mathbf{s}$ be the mapping from $\mathbb{R}^N$ into $\mathbb{C}^N$ defined by $s_n = \sqrt{p_n}\exp(j\phi_n)$ where $\mathbf{p} \in \mathbb{R}_+^N$ and $\boldsymbol{\phi} \in \mathbb{R}^N$. For convenience, we suppress the dependence of $\mathbf{s}$ on $\boldsymbol{\phi}$. Then,

$$\frac{\partial s_n}{\partial \phi_p} = \begin{cases} js_p & n = p \\ 0 & n \neq p \end{cases} \text{ and } \frac{\partial s_n^*}{\partial \phi_p} = \begin{cases} -js_p^* & n = p \\ 0 & n \neq p \end{cases} \tag{C.30}$$

For any composition of the form $\widehat{G} = G \circ \mathbf{s}$ with $G : \mathbb{C}^N \to \mathbb{R}$, the conjugate symmetry in (C.30) and the conjugate symmetry between (4) and (5) for real-valued functions of real variables can be used to find the partial of $\widehat{G}$ with respect to ϕ_p. Beginning with an

application of the multivariate chain rule, we have

$$\frac{\partial \widehat{G}}{\partial \phi_p} = \sum_{n=1}^{N} \left\{ \frac{\partial G}{\partial s_n}\frac{ds_n}{d\phi_p} + \frac{\partial G}{\partial s_n^*}\frac{ds_n^*}{d\phi_p} \right\} \tag{C.31}$$

$$= \sum_{n=1}^{N} \left\{ \frac{\partial G}{\partial s_n}\frac{ds_n}{d\phi_p} + \left[\frac{\partial G}{\partial s_n}\frac{ds_n}{d\phi_p}\right]^* \right\} \tag{C.32}$$

$$= \sum_{n=1}^{N} 2\mathrm{Re}\left\{ \frac{\partial G}{\partial s_n}\frac{ds_n}{d\phi_p} \right\} \tag{C.33}$$

$$= 2\mathrm{Re}\left\{ \frac{\partial G}{\partial s_p}\frac{ds_p}{d\phi_p} \right\} \tag{C.34}$$

$$= -2\mathrm{Im}\left\{ s_p \frac{\partial G}{\partial s_p} \right\} \tag{C.35}$$

This implies

$$\nabla_{\boldsymbol{\phi}} \widehat{G} = -2\mathrm{Im}\left\{ \mathbf{s}^T \odot \nabla_{\mathbf{s}} G \right\} \tag{C.36}$$

Applying result (C.24) to (C.36) yields

$$\nabla_{\boldsymbol{\phi}} \widehat{\mathcal{S}} = -2\mathrm{Im}\left\{ \mathbf{s}^T \odot \nabla_{\mathbf{s}} S \right\} \tag{C.37}$$

$$= 2\mathrm{Im}\left\{ \mathbf{s}^T \odot (\mathbf{s}_d - \mathbf{s})^H \right\} \tag{C.38}$$

C.3 SINR Gradient for NLFM

The gradient of the cost function (4.15) with respect to the phase parameters (Fourier coefficients) can be efficiently calculated as follows. First, note that

$$\frac{d}{da_p} s(n) = \frac{d}{da_p} \exp\left(j\left(\phi_n + \sum_{k=0}^{M} a_k \sin(\omega_k n) + b_k \cos(\omega_k) \right) \right) \tag{C.39}$$

$$= j\sin(\omega_p n)\, s(n) \tag{C.40}$$

and define the vector $\mathbf{x}^{(p)}$ by $x_m^{(p)} = \sin(\omega_p m)$. Then,

$$\frac{d}{da_p} J(\mathbf{s}) = \frac{d}{da_p} \mathbf{s}^H \mathbf{K} \mathbf{s} \tag{C.41}$$

$$= \frac{d}{da_p} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} [\mathbf{K}]_{n,m} s_n^* s_m \tag{C.42}$$

$$= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} [\mathbf{K}]_{n,m} \left(s_n^* \frac{d}{da_p} s_m + s_m \frac{d}{da_p} s_n^* \right) \tag{C.43}$$

$$= j \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} [\mathbf{K}]_{n,m} \left(s_n^* s_m \sin(\omega_p m) - s_m s_n^* \sin(\omega_p n) \right) \tag{C.44}$$

$$= j \mathbf{s}^H \mathbf{K} \left(\mathbf{x}^{(p)} \odot \mathbf{s} \right) - j \left(\mathbf{x}^{(p)} \odot \mathbf{s} \right)^H \mathbf{K} \mathbf{s} \tag{C.45}$$

$$= 2\mathrm{Re} \left\{ -j \left(\mathbf{x}^{(p)} \odot \mathbf{s} \right)^H \mathbf{K} \mathbf{s} \right\} \tag{C.46}$$

$$= 2\mathrm{Im} \left\{ \left(\mathbf{x}^{(p)} \odot \mathbf{s} \right)^H \mathbf{K} \mathbf{s} \right\} \tag{C.47}$$

And therefore,

$$\nabla_{\mathbf{a}} J(\mathbf{s}) = 2\mathrm{Im} \left\{ \mathbf{S} \left(\mathbf{s}^* \odot \mathbf{K} \mathbf{s} \right) \right\} \tag{C.48}$$

where $[\mathbf{S}]_{n,m} = \sin(\pi(m-1)(n-1)/M)$. Similarly,

$$\nabla_{\mathbf{b}} J(\mathbf{s}) = 2\mathrm{Im} \left\{ \mathbf{C} \left(\mathbf{s}^* \odot \mathbf{K} \mathbf{s} \right) \right\} \tag{C.49}$$

where $[\mathbf{C}]_{n,m} = \cos(\pi(m-1)(n-1)/M)$.

Bibliography

[1] A. Sayed, *Fundamentals of Adaptive Filtering*. Wiley-Interscience, 2003.

[2] A. van den Bos, "Complex gradient and Hessian," *Vision, Image and Signal Processing, IEE Proceedings* -, vol. 141, no. 6, pp. 380–383, Dec. 1994.

[3] [Online]. Available: http://www.waveformdiversity.org/

[4] [Online]. Available: http://signal.ese.wustl.edu/MURI/index.html

[5] [Online]. Available: http://signal.ese.wustl.edu/DARPA/index.html

[6] *Journal of Selected Topics in Signal Processing, IEEE*, vol. 1, no. 1, Jun. 2007.

[7] *Signal Processing Magazine, IEEE*, vol. 26, no. 1. IEEE, Jan. 2009.

[8] B. Moran, , S. Suvorova, and S. Howard, *Sensor Management for Radar: A Tutorial*, ser. NATO Security through Science Series. Springer Netherlands, 2006, vol. 2, pp. 269–291.

[9] Sandeep P. Sira, Antonia Papandreou Suppappola, Darryl Morrell, and Douglas Cochran, "Waveform-Agile Tracking In Heavy Sea Clutter," in *International Waveform Diversity & Design Conference*, Jun. 4–8, 2007.

[10] H. L. Van Trees, "Optimum Signal Design and Processing for Reverberation-Limited Environments," *Military Electronics, IEEE Transactions on*, vol. 9, no. 3, pp. 212–229, July 1965.

[11] D. DeLong and E. Hofstetter, "On the design of optimum radar waveforms for clutter rejection," *IEEE Trans. Inf. Theory*, vol. 13, no. 3, pp. 454–463, Jul. 1967.

[12] ——, "Optimum Radar Signal-Filter Pairs in a Cluttered Environment," *IEEE Trans. Inf. Theory*, vol. 16, no. 1, pp. 89–90, Jan. 1970.

[13] W. D. Rummler, "A technique for improving the clutter performance of coherent pulse train signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 3, pp. 898–906, Nov. 1967.

[14] M. Ares, "Optimum burst waveforms for detection of targets in uniform range-extended clutter," *IEEE Trans. Aerospace and Electronic Systems*, vol. 3, p. 138, ,January 1967.

[15] J. S. Thompson and E. L. Titlebaum, "The design of optimal radar waveforms for clutter rejection using the maximum principle," *Supplement to IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-3, pp. 581–589, Nov. 1967.

[16] L. Spafford, "Optimum radar signal processing in clutter," *IEEE Trans. Inf. Theory*, vol. 14, no. 5, pp. 734–743, Sep. 1968.

[17] Thomas G. Kincaid, "Optimum Waveforms for Correlation Detection in the Sonar Environment: Noise-Limited Conditions," *The Journal of the Acoustical Society of America*, vol. 43, no. 2, pp. 258–268, 1968.

[18] ——, "On Optimum Waveforms for Correlation Detection in the Sonar Environment: Reverberation-Limited Conditions," *The Journal of the Acoustical Society of America*, vol. 44, no. 3, pp. 787–796, 1968.

[19] D. DeLong and E. Hofstetter, "Correction to 'The design of clutter-resistant radar waveforms with limited dynamic range'," *IEEE Trans. Inf. Theory*, vol. 16, no. 1, pp. 73–73, Jan. 1970.

[20] G. W. Zeoli, "Some Results on Pulse-Burst Radar Design," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 7, no. 3, pp. 486–498, May 1971.

[21] S. Lee and J. Uhran, J., "Optimum signal and filter design in underwater acoustic echo ranging system," in *OCEANS*, vol. 4, Sep. 1972, pp. 25–30.

[22] S. P. Lee and J. J. Uhran, "Optimum Signal and Filter Design in Underwater Acoustic Echo Ranging Systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 9, no. 5, pp. 701–713, Sep. 1973.

[23] M. F. Mesiya and P. J. Mclane, "Design of Optimal Radar Signals Subject to a Fixed Amplitude Constraint," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 9, no. 5, pp. 679–687, Sep. 1973.

[24] A. Cohen, "A nonlinear integer programming algorithm for the design of radar waveforms," *Annual Allerton Conference on Circuit and System Theory, 12th, Monticello, Ill., October 2-4, 1974,*, 1975.

[25] A. I. Cohen, "An Algorithm for Designing Burst Waveforms with Quantized Transmitter Weights," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 11, no. 1, pp. 56–64, Jan. 1975.

[26] L. Sibul and E. Titlebaum, "Signal design for detection of targets in clutter," *Proceedings of the IEEE*, vol. 69, no. 4, pp. 481–482, April 1981.

[27] S. M. Kay and J. H. Thanos, "Optimal transmit signal design for active sonar/radar," in *Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on*, vol. 2, Orlando, FL, USA, 2002, pp. 1513–1516.

[28] S. Kay, "Optimal Signal Design for Detection of Gaussian Point Targets in Stationary Gaussian Clutter/Reverberation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 31–41, Jun. 2007.

[29] T. Kooij, "Optimum signals in noise and reverberation," in *NATO Advanced Study Institute on Signal Processing With Emphasis on Underwater Acoustics*, vol. I, Enschede, The Netherlands, 1968.

[30] Ric A. Romero, Jun Hyeong Bae, and Nathan A. Goodman, "Theory and Application of SNR- and MI-Based Matched Illumination Waveforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. Pre-Print, 2009.

[31] M. R. Bell, "Information theory and radar waveform design," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1578–1597, Sep. 1993.

[32] J. S. Bergin, P. M. Techau, J. E. Don Carlos, and J. R. Guerci, "Radar waveform optimization for colored noise mitigation," in *Radar Conference, 2005 IEEE International*, May 9–12, 2005, pp. 149–154.

[33] J. Li, J. R. Guerci, and L. Xu, "Signal Waveform's Optimal-under-Restriction Design for Active Sensing," *IEEE Signal Process. Lett.*, vol. 13, no. 9, pp. 565–568, Sep. 2006.

[34] A. De Maio and A. Farina, "Code Selection for Radar Performance Optimization," in *International Waveform Diversity & Design Conference*, Jun. 4–8, 2007.

[35] S. De Nicola, Y. Huang, S. De Maio, A.and Zhang, and A. Farina, "Code Optimization with Similarity and Accuracy Constraints," in *Radar Conference, 2008 IEEE*, May 26–30, 2008.

[36] A. De Maio, S. De Nicola, Yongwei Huang, Shuzhong Zhang, and A. Farina, "Code Design to Optimize Radar Detection Performance Under Accuracy and Similarity Constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5618–5629, Nov. 2008.

[37] M. J. Lindenfeld, "Sparse frequency transmit-and-receive waveform design," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 3, pp. 851–861, Jul. 2004.

[38] L. Patton and B. Rigling, "Nonquadratic regularization for waveform optimization," in *Radar, 2006 IEEE Conference on*, Apr. 24–27, 2006.

[39] M. Athans and F. C. Schweppe, "Optimal waveform design via control theoretic principles," *Information Control*, vol. 10, pp. 335–377, 1967.

[40] D. DeLong and E. Hofstetter, "The design of clutter-resistant radar waveforms with limited dynamic range," *IEEE Trans. Inf. Theory*, vol. 15, no. 3, pp. 376–385, May 1969.

[41] D. DeLong, "Design of radar signals and receivers subject to implementation errors," *IEEE Trans. Inf. Theory*, vol. 16, no. 6, pp. 707–711, Nov. 1970.

[42] H. Van Trees, *Detection, estimation, and modulation theory Part III*. Wiley, 2001.

[43] A. Balakrishnan, "Signal design for a class of clutter channels (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 170–173, Jan. 1968.

[44] J. C. Holtzman and J. S. Thorp, "Optimum Signals for Radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 5, no. 6, pp. 898–905, Nov. 1969.

[45] K. T. Wong, "Adaptive pulse-diverse radar/sonar waveform design," in *Radar Conference, 1998. RADARCON 98. Proceedings of the 1998 IEEE*, Dallas, TX, USA, May 11–14, 1998, pp. 105–110.

[46] K. T. Wong and Wing-Kit Chung, "Pulse-diverse radar/sonar FSK-PSK waveform design to emphasize/de-emphasize designated Doppler-delay sectors," in *Radar Conference, 2000. The Record of the IEEE 2000 International*, Alexandria, VA, May 7–12, 2000, pp. 745–749.

[47] Wing-Kit Chung and Kainam Thomas Wong, "Pulse-diverse radar waveform design for accurate joint estimation of time delay and Doppler shift," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, vol. 5, Istanbul, Turkey, 2000, pp. 3037–3040.

[48] R. J. Bonneau and M. C. Wicks, "A numerical waveform design approach to decorrelate target and noise," in *Radar Conference, 2001. Proceedings of the 2001 IEEE*, Atlanta, GA, May 1–3, 2001, pp. 448–450.

[49] K. Vastola, J. Farnbach, and S. Schwartz, "Maximin sonar system design for detection," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '83.*, vol. 8, Apr. 1983, pp. 591–594.

[50] S. P. Sira, D. Cochran, A. Papandreou Suppappola, D. Morrell, W. Moran, S. D. Howard, and R. Calderbank, "Adaptive Waveform Design for Improved Detection of Low-RCS Targets in Heavy Sea Clutter," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 56–66, Jun. 2007.

[51] K. D. Ward, C. J. Baker, and S. Watts, "Maritime surveillance radar. I. Radar scattering from the oceansurface," *Radar and Signal Processing, IEE Proceedings F*, vol. 137, no. 2, pp. 51–62, Apr. 1990.

[52] K. J. Sangston and K. R. Gerlach, "Coherent detection of radar targets in a non-gaussian background," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 2, pp. 330–340, Apr. 1994.

[53] C. J. Oliver, "Correlated k-distributed clutter models," *Journal of Modern Optics*, vol. 32, no. 12, pp. 1515–1547, 1985.

[54] A. Farina, F. Gini, M. V. Greco, and L. Verrazzani, "High resolution sea clutter data: statistical analysis of recordedlive data," *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 144, no. 3, pp. 121–130, Jun. 1997.

[55] M. Skolnik, *Introduction to radar systems*, 3rd ed. McGraw-Hill New York, 2001.

[56] D. Gjessing, *Target Adaptive Matched Illumination Radar: Principles & Applications*. Peter Peregrinus on behalf of the Institution of Electrical Engineers, 1986.

[57] C. Pell, "Book review: Target Adaptive Matched Illumination Radar: Principles and Applications," *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 133, no. 6, pp. 581–582, October 1986.

[58] E. Kennaugh, "The K-pulse concept," *IEEE Trans. Antennas Propag.*, vol. 29, no. 2, pp. 327–331, Mar. 1981.

[59] C. E. Baum, E. J. Rothwell, K. M. Chen, and D. P. Nyquist, "The singularity expansion method and its application to targetidentification," *Proc. IEEE*, vol. 79, no. 10, pp. 1481–1492, Oct. 1991.

[60] E. Jaynes, "Theory of radar target discrimination," Tech. Rep. RD-AS-96-6, Procurement Instrument Identification Number DAAL03-86-D-001, DO 1515, Redstone Arsenal, AL, Feb. 1991, Tech. Rep., 1991.

[61] Lloyd S. Riggs, Michael A. Richards, Thomas H. Shumpert, and C. Ray Smith, "On waveform design for optimal target discrimination," *International Journal of Imaging Systems and Technology*, vol. 4, pp. 327–335, 1992.

[62] H. H. Schreiber and M. G. O'Connor, "Adaptive waveform radar," United States Patent 4,901,082, Feb. 1990.

[63] P. G. Grieve and J. R. Guerci, "Optimum Matched Illumination-Reception Radar," United States Patent 5,175,552, Dec. 1992.

[64] J. R. Guerci and P. G. Grieve, "Optimum Matched Illumination Waveform Design Process," United States Patent 5,121,125, Jun. 1992.

[65] J. R. Guerci, R. W. Schutz, and J. D. Hulsmann, "Constrained Optimum Matched Illumination-Reception Radar," United States Patent 5,146,229, Sep. 1992.

[66] J. R. Guerci, "Optimum matched illumination-reception radar for target classification," United States Patent 5,381,154, Jan. 1995.

[67] ——, "Method of recognizing a radar target object type and apparatus therefor," United States Patent 5,392,050, Feb. 1995.

[68] T. W. Barrett, "Active Signaling Systems," United States Patent 5,486,833, Jun. 1996.

[69] R. C. Poe and C. E. Hein, "Clutter normalization by adaptation of transmit waveform," United States Patent 6,861,974 B1, Mar. 2005.

[70] S. U. Pillai, "Target identification from a pool of targets using a new adaptive transmitter-receiver design," United States Patent 7,038,616, May 2006.

[71] ——, "Target identification from a pool of targets using a new adaptive transmitter-receiver design," United States Patent 7,132,977, Nov. 2006.

[72] W. Adams, Vinh; Dwelly and R. J. Adams, "Radar System and Method for Reducing Clutter in a High-Clutter Environment," United States Patent 7,236,124 B2, Jun. 2007.

[73] V. Adams and W. Dwelly, "Radar System with Adaptive Waveform Processing and Methods for Adaptively Controlling the Shape of a Radar Ambiguity Function," United States Patent 7,327,307 B2, Feb. 2008.

[74] M. L. Honig, K. Steiglitz, and S. A. Norman, "Optimization of signal sets for partial-response channels. I. Numerical techniques," *IEEE Trans. Inf. Theory*, vol. 37, no. 5, pp. 1327–1341, Sep. 1991.

[75] M. L. Honig, K. Steiglitz, V. Balakrishnan, and E. Rantapaa, "Discrete-time signal design for maximizing separation in amplitude," *IEEE Trans. Inf. Theory*, vol. 41, no. 1, pp. 164–170, Jan. 1995.

[76] S. Sowelam and A. H. Tewfik, "Optimal waveform selection in range-Doppler imaging," in *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, vol. 1, Austin, TX, USA, Nov. 13–16, 1994, pp. 441–445.

[77] D. T. Gjessing, "Matched illumination radar involving synthetic aperture andsynthetic pulse for signal to clutter enhancement," in *Radar, 1996. Proceedings., CIE International Conference of*, Beijing, China, Oct. 8–10, 1996, pp. 594–600.

[78] S. U. Pillai, D. C. Youla, H. S. Oh, and J. R. Guerci, "Optimum transmit-receiver design in the presence ofsignal-dependent interference and channel noise," in *Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on*, vol. 2, Pacific Grove, CA, USA, 1999, pp. 870–875.

[79] S. U. Pillai, H. S. Oh, D. C. Youla, and J. R. Guerci, "Optimal transmit-receiver design in the presence of signal-dependent interference and channel noise," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 577–584, Mar. 2000.

[80] S. U. Pillai, K. Y. Li, and B. Himed, *Space Based Radar: Theory & Applications*. McGraw-Hill, 2007.

[81] J. R. Guerci and S. U. Pillai, "Theory and Application of Adaptive Transmission (ATx) Radar," 2000.

[82] ——, "Theory and application of optimum transmit-receive radar," in *Radar Conference, 2000. The Record of the IEEE 2000 International*, Alexandria, VA, USA, 2000, pp. 705–710.

[83] ——, "Adaptive transmission radar: the next "wave"?" in *National Aerospace and Electronics Conference, 2000. NAECON 2000. Proceedings of the IEEE 2000*, Dayton, OH, USA, 2000, pp. 779–786.

[84] D. A. Garren, M. K. Osborn, A. C. Odom, J. S. Goldstein, S. Unnikrishna Pillai, and J. R. Guerci, "Optimization of single transmit pulse shape to maximize detection and identification of ground mobile targets," in *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, vol. 2, Pacific Grove, CA, Oct. 29–Nov. 1, 2000, pp. 1535–1539.

[85] D. A. Garren, M. K. Osborn, A. C. Odom, J. S. Goldstein, S. U. Pillai, and J. R. Guerci, "Optimal transmission pulse shape for detection and identification with uncertain target aspect," in *Radar Conference, 2001. Proceedings of the 2001 IEEE*, Atlanta, GA, May 1–3, 2001, pp. 123–128.

[86] ——, "Enhanced target detection and identification via optimised radar transmission pulse shape," *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 148, no. 3, pp. 130–138, Jun. 2001.

[87] S. U. Pillai, H. S. Oh, and J. R. Guerci, "Multichannel matched transmit-receiver design in presence ofsignal-dependent interference and noise," in *Sensor Array and Multichannel Signal Processing Workshop. 2000. Proceedings of the 2000 IEEE*, Cambridge, MA, USA, 2000, pp. 385–389.

[88] D. A. Garren, A. C. Odom, M. K. Osborn, J. S. Goldstein, S. U. Pillai, and J. R. Guerci, "Full-polarization matched-illumination for target detection and identification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 824–837, Jul. 2002.

[89] S. U. Pillai, J. R. Guerci, and S. R. Pillai, "Joint optimal Tx-Rx design for multiple target identification problem," in *Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002*, Aug. 4–6, 2002, pp. 553–556.

[90] S. U. Pillai and H. S. Oh, "Optimum MIMO transmit-receiver design in presence of interference," in *Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on*, vol. 4, May 25–28, 2003.

[91] S. Haykin and B. Currie, "Literature Search on Adaptive Radar Transmit Waveforms," *ASL Report*, pp. 02–01, 2002.

[92] T. Kirubarajan and S. Haykin, "Simulation Results on Adaptive Radar Transmit Waveforms (Final Report)," McMaster University, Tech. Rep., 2003.

[93] S. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Prentice Hall PTR, 1998.

[94] Hao Tan and Min Zheng, "Transmit signal design for optimal deconvolution," *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 2, pp. 553–6, May 17–21, 2004.

[95] S. M. Sowelam and A. H. Tewfik, "Waveform selection in radar target classification," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 1014–1029, May 2000.

[96] P. Venkata and N. Goodman, "Novel iterative techniques for radar target discrimination," in *2006 International Waveform Diversity and Design Conference*, Lihue, HI, Jan. 2006.

[97] Jun Hyeong Bae and Nathan A. Goodman, "Adaptive Waveforms for Target Class Discrimination," in *International Waveform Diversity & Design Conference*, Jun. 4–8, 2007.

[98] N. A. Goodman, P. R. Venkata, and M. A. Neifeld, "Adaptive Waveform Design and Sequential Hypothesis Testing for Target Recognition With Active Sensors," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 105–113, Jun. 2007.

[99] T. B. Butler and N. A. Goodman, "Multistatic target classification with adaptive waveforms," in *Radar Conference, 2008 IEEE*, May 26–30, 2008.

[100] Thomas Benjamin Butler, "Multistatic Target Classificaiton With Adaptive Waveforms," Master's thesis, The University of Arizona, 2008.

[101] P. Woodward, *Probability and Information Theory, with Applications to Radar*. London: Pergamon Press, 1953.

[102] N. Levanon and E. Mozeson, *Radar Signals*. Wiley-IEEE Press, 2004.

[103] M. Richards, *Fundamentals of Radar Signal Processing*. McGraw-Hill, 2005.

[104] L. K. Patton and B. D. Rigling, "Modulus Constraints in Adaptive Radar Waveform Design," in *Radar, 2008 IEEE Conference on*, Feb. 8–13 2008.

[105] L. K. Patton and B. D. Rigling, "Alternating projections for envelope constraint satisfaction in radar waveform optimization," *IEEE Trans. Aerosp. Electron. Syst.*, May 2009, submitted.

[106] ——, "On the structure of radar waveform optimization problems," *IEEE Trans. Aerosp. Electron. Syst.*, Jun. 2009, submitted.

[107] L. K. Patton and B. D. Rigling, "Autocorrelation and Modulus Constraints in Radar Waveform Optimization," in *4th Int'l Waveform Diversity and Design Conference*, Feb. 8–13 2009.

[108] L. K. Patton and B. D. Rigling, "Efficient constrainted design of constant modulus waveforms," *IET Radar Sonar & Navigation*, Jun. 2009, submitted.

[109] Merrill Skolnik, Ed., *Radar Handbook*, 2nd ed. McGraw-Hill, 1990.

[110] T. Varslot, C. E. Yarman, M. Cheney, and B. Yazici, "A variational approach to waveform design for synthetic aperture imaging," *Inverse Problems and Imaging*, vol. 1, no. 3, pp. 577–592, 2007.

[111] S. Haykin, "Optimal Waveform Design for Cognitive Radar," in *Signals, Systems and Computers, 2008. Conference Record of the Forty-Second Asilomar Conference on*, vol. 1, Pacific Grove, CA, USA, 2008, presentation Slides.

[112] A. Leshem, O. Naparstek, and A. Nehorai, "Information Theoretic Adaptive Radar Waveform Design for Multiple Extended Targets," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 42–55, Jun. 2007.

[113] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, pp. 2758–2769, 1982.

[114] H. H. Bauschke, P. C. Combettes, and D. R. Luke, "On the structure of Some Phase Retrieval Algorithms," in *Proceedings of the IEEE International Conference on Image Processing*, vol. II, Rochester, NY, September 22-25 2002, pp. 841–844.

[115] H. H. Bauschke, , P. L. Combettes, and D. R. Luke, "A New Generation of Iterative Transform Algorithms for Phase Contrast Tomography," in *Proceedings of the IEEE 2005 International Conference on Acoustics, Speech and Signal Processing,*, 2005.

[116] H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization," *J. Opt. Soc. Am. A*, vol. 19, no. 7, pp. 1334–1345, 2002.

[117] D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse Problems*, vol. 21, pp. 37–50, 2005.

[118] A. S. Lewis, D. R. Luke, and J. Malick, "Local convergence for alternating and averaged nonconvex projections," *Foundations of Computational Mathematics*, vol. (to appear), 2008.

[119] D. R. Luke, "Finding Best Approximation Pairs Relative to a Convex and Prox-Regular Set in a Hilbert Space," *SIAM J. on Optimization*, vol. 19, no. 2, pp. 714–739, 2008.

[120] J. Fienup, "Iterative Method Applied to Image Reconstruction and to Computer Generated Holograms," *Optical Engineering*, vol. 19, pp. 297–305, May-June 1980 1980.

[121] A. Rundquist, A. Efimov, and D. H. Reitze, "Pulse shaping with the Gerchberg-Saxton algorithm," *J. Opt. Soc. Am. B*, vol. 19, no. 10, pp. 2468–2478, 2002.

[122] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

[123] D. C. Youla and H. Webb, "Image Restoration by the Method of Convex Projections: Part 1Theory," *Medical Imaging, IEEE Transactions on*, vol. 1, no. 2, pp. 81–94, Oct. 1982.

[124] A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," *J. Opt. Soc. Am.*, vol. A 1, pp. 932–943, 1984.

[125] V. Elser, "Phase retrieval by iterated projections," *J. Opt. Soc. Am.*, vol. A 20, pp. 40–55, 2003.

[126] M. Hacker, G. Stobrawa, and T. Feurer, "Iterative Fourier transform algorithm for phase-only pulse shaping," *Opt. Express*, vol. 9, pp. 191–199, 2001.

[127] Leland Jackson, Steven Kay, and Naresh Vankayalapati, "Nonlinear FM Synthesis for Radar," sent to me by Kay in June 2008.

[128] D. Fuhrmann, P. Browning, and M. Rangaswamy, "Constant-modulus partially correlated signal design for uniform lienar & rectangular mimo radar arrays," in *Proceedings of the 4th International Waveform Diversity & Design Conference*, Orlando, FL, Feb. 2009.

[129] S. U. Pillai, K. Y. Li, and H. Beyer, "Reconstruction of Constant Envelop Signals with Given Fourier Transform Magnitude," *Proceedings of the IEEE Radar Conference*, 2009.

[130] J. Arora, *Introduction to Optimum Design*, 2nd ed. New York: McGraw-Hill, 2004.

[131] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. Springer, 2006.

[132] S. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993.

[133] E. C. Farnett and G. H. Stevens, "Pulse compression radar," in *Radar Handbook*, 2nd ed., M. Skolnik, Ed. McGraw-Hill, 1990, ch. 10, pp. 10.1–10.39.

[134] M. Soumekh, *Synthetic aperture radar signal processing*. Wiley New York, NY, 1999.

[135] J. W. Taylor, Jr., "Receivers," in *Radar Handbook*, 2nd ed., M. Skolnik, Ed. McGraw-Hill, 1990, ch. 3, pp. 3.1–3.46.

[136] *Optimization Toolbox 4: User's Guide*, 2008. [Online]. Available: http://www.mathworks.com/access/helpdesk/help/pdf_doc/optim/optim_tb.pdf

[137] P. Stoica and R. Moses, *Spectral analysis of signals*. Pearson Prentice Hall, 2005.

[138] R. Gray, "Toeplitz and Circulant Matrices: A Review," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[139] S. Sussman, "Least-square synthesis of radar ambiguity functions," *IEEE Trans. Inf. Theory*, vol. 8, no. 3, pp. 246–254, Apr. 1962.

[140] B. P. Lathi, *Signal Processing and Linear Systems*. Berkely-Cambridge Press, 1998.